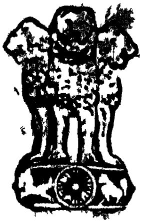


सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

I.A.R.I.6.
GIP NLK—H-3 I.A.R.I.—10-5 55—15,000

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A. MATHEMATICAL AND PHYSICAL SCIENCES

VOL 186

LONDON
Printed and published for the Royal Society
By the Cambridge University Press
Bentley House, N.W.1

24 September 1946

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
Cambridge, and Bentley House, London

Agents for Canada and India: Macmillan

CONTENTS

SERIES A VOLUME 186

No. A 1004. 4 June 1946

PAGE

Bakerian Lecture. The structure, function and synthesis of polysaccharides. By W. N. Haworth, F.R.S. 1

Turbulence and diffusion in the lower atmosphere. By R. Frost 20

Contributions to the theory of chromatography. By E. Glückauf . . . 35

The plastic deformation of non-cubic metals by heating and cooling. By W. Boas and R. W. K. Honeycombe. (Plates 1–5) 57

A contribution to the theory of the Foucault test. By E. H. Linfoot . . . 72

An integral-equation method of solving plane diffraction problems. By E. T. Copson 100

On the divergence difficulty of quantized field theories and the rigorous treatment of radiation reaction. By H. W. Peng 119

No. A 1005. 9 July 1946

A discussion on units and standards. By Sir Charles Darwin, F.R.S., and others. (Plate 6) 149

Kinetic studies in the chemistry of rubber and related materials. I. The thermal oxidation of ethyl linoleate. By J. L. Bolland 218

Recombination and the long duration of the Balmer spectrum. By H. Zanstra . 236

The emission of light from spark discharges. By J. D. Craggs and J. M. Meek. (Plates 7–11) 241

Further interferometric studies with mica. New multiple-beam fringes and their application. By S. Tolansky. (Plates 12–14) 261

No. A 1006. 24 September 1946

PAGE

The air wave surrounding an expanding sphere. By G. I. Taylor, F.R.S. . . . 273

Explosion waves and shock waves. VI. The disturbance produced by bursting diaphragms with compressed air. By W. Payman and W. C. F. Shepherd. (Plates 15–18) 293

The diffraction of sound pulses. I. Diffraction by a semi-infinite plane. By F. G. Friedlander 322

The diffraction of sound pulses. II. Diffraction by an infinite wedge. By F. G. Friedlander 344

The diffraction of sound pulses. III. Note on an integral occurring in the theory of diffraction by a semi-infinite screen. By F. G. Friedlander 352

The diffraction of sound pulses. IV. On a paradox in the theory of reflexion. By F. G. Friedlander 356

The rate of evaporation of droplets. Evaporation and diffusion coefficients, and vapour pressures of dibutyl phthalate and butyl stearate. By R. S. Bradley, M. G. Evans and R. W. Whytlaw-Gray, F.R.S. 368

Fluid motion between parallel planes. Dynamical stability. By D. Meksyn . 391

No. A 1007. 24 September 1946

Experimental production of red and yellow sandstones from chalybeate water. By Lord Rayleigh, F.R.S. 411

The probability of encounters between gas molecules. By Lewis F. Richardson, F.R.S. 422

On the conservation of momentum. By E. A. Milne, F.R.S. 432

Kinetics of the base-catalysed bromination of diethyl malonate. By R. P. Bell, F.R.S., D. H. Everett and H. C. Longuet-Higgins 443

An invariant form for the prior probability in estimation problems. By Harold Jeffreys, F.R.S. 453

The mechanism of the hydrogen-oxygen reaction. IV. The activation energy of the initiating process. By C. F. Cullis and C. N. Hinshelwood, F.R.S. . . 462

The mechanism of the hydrogen-oxygen reaction. V. The reaction in vessels coated with alkali iodides. By C. F. Cullis and C. N. Hinshelwood, F.R.S. . . . 469

The drag of cylinders in fluids at slow speeds. By C. M. White 472

The theory of axisymmetric turbulence. By G. K. Batchelor 480

The correspondence between the particle and the wave aspects of the meson and the photon. By Harish-Chandra 502

Index 527

BAKERIAN LECTURE

The structure, function and synthesis of polysaccharides

By W. N. HAWORTH, F.R.S.

(*Received* 31 *July* 1945. *Delivered* 15 *June* 1944)

Cellulose is probably the most abundant organic compound occurring in nature, where it plays an all-important role in building up the skeletal framework of the plant cell wall. The accepted molecular structure of cellulose represented as repeating units of cellobiose was derived by chemical methods which I have previously summarized (Haworth 1939), and the constitution so assigned has received support in a remarkable way by the X-ray studies of Sponsler and Dore and of Meyer and Mark. The skeleton chain of cellulose can be illustrated by the model shown in figure 1, where four β-glucopyranose units are assembled side by side, such an arrangement being repeated many times to give a long extended chain structure. The arrangement is perfectly symmetrical since the hexagons are linked through the 1:4-positions, and the side chains, representing the primary alcohol groups at C_6 of the glucose residues, lie alternately above and below the axis of the main chain. In this representation each alternate hexose is rotated through 180° so that deviation from the linear structure is almost completely obviated. The constitution is in keeping with the fact that cellulose occurs in fibrous form and is highly resistant to solvent action and to chemical attack. It is suggested that there are possibly some infrequent points of bonding between adjoining chains to provide a kind of cross-linkage.

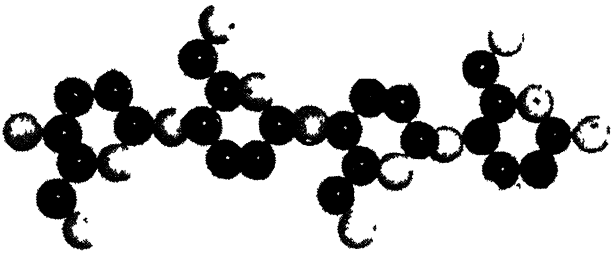


FIGURE 1

It will be apparent from the above model that the constitution of cellobiose is the important factor in deciding the structure of cellulose, and this crystalline disaccharide can readily be obtained from it by acetolysis methods. It has been my endeavour to support the constitutional studies of this fundamental repeating unit by simple direct synthesis, a result which has now been attained by three of my colleagues (Stacey, Smith & Gilbert 1944). The sodio-derivative of 1:2:3:6-tetra-acetyl glucopyranose was condensed with α-acetobromo-glucose (figure 2) in the liquid state, thereby forming crystalline octa-acetyl cellobiose which could be smoothly de-acetylated to cellobiose. An earlier and ingenious synthesis of cellobiose had been reported (Haskins, Hann & Hudson 1942); but that which I have outlined above is the more direct, and both syntheses furnish a convincing proof of its constitution and of its recognition as glucopyranose-1:4-β-glucopyranose. A notable synthesis of cellulose from glucose, using the micro-organism *Acetobacter Xylinum*, has been achieved by Hibbert (1931).

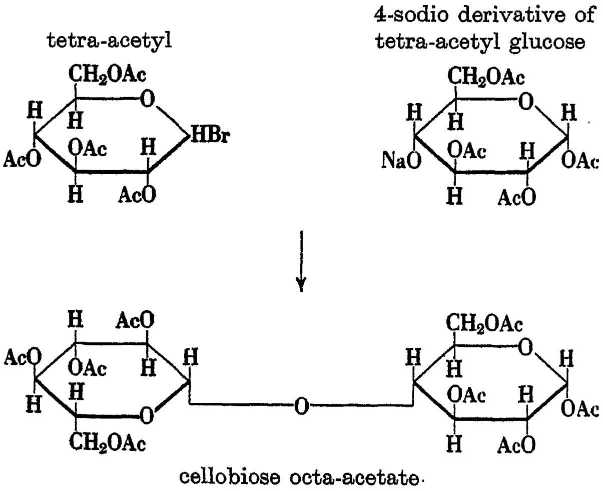


FIGURE 2

Starch, which functions as a food reserve in plants, and glycogen, an energy source in animals, contain the disaccharide maltose as a repeating unit corresponding to the repeating unit of cellobiose in cellulose. Maltose and cellobiose are identical in structure but not in configuration, maltose being glucopyranose-1:4-α-glucopyranose. Natural starch, prepared from a variety of sources, contains two different polysaccharides, namely, amylose (20 %) which gives a deep blue coloration with iodine, and amylopectin (80 %), giving a reddish coloration under the same conditions of test. Various methods are available for the separation of these two components, a suitable one being the butyl alcohol fractionation devised by Schoch (1942); but new and less complex methods of separation have recently been devised by my co-workers. It is now, indeed, a comparatively simple procedure to obtain either component in the pure state. As already described, amylose is a long-chain component (Hassid & McCready 1943), while amylopectin consists of highly aggregated or branched short chains, a comparison of their properties being shown in the table (figure 3) from a publication by R. W. Kerr.

	straight-chain amylose	branched-chain amylopectin
solubility in H_2O	originally soluble as isolated	soluble after separation from amylose
colloidal stability	spontaneously precipitates from solution $> 0{\cdot}5$ %	very stable in 5 % solutions
gel properties	5 % solution sets quickly to tough, irreversible gels	form soft gels which redissolve
viscosity (dilute solution)	lower, but enormous increase with concentration	higher, simple function of concentration
protective colloid effect (clay)	low, negative at high concentration	high
iodine colour	blue	reddish purple
enzymes:		
(1) liquefying (α-amylases)	attacked very slowly. Retrogrades and is then resistant	readily converted to lower dextrins
(2) saccharifying (β-amylase)	readily converted 95–100 % to maltose at low substrate concentration. At high concentration retrogrades to resistant form	converts 52 % to maltose, 48 % to soluble dextrin (dextrin-A)
estimated mol.wt	35,000 (200 glucose units)	350,000 (2000 glucose units)

FIGURE 3. Comparison of corn amylopectin and amylose. (From Kerr 1942.)

An outstanding achievement of recent years has been the study of the enzyme systems concerned in the orderly breakdown of starch to glucose-1-phosphate and the resynthesis of amylose from this product by the same agency (Hanes 1940). This has revealed the mechanism of the biological synthesis of a starch polysaccharide from glucose. Synthetic amyloses produced from glucose-1-phosphate by the potato phosphorylase of Hanes and by the muscle phosphorylase of Cori & Cori (1939) have properties remarkably similar to those of natural amylose. They are completely hydrolysed to maltose by the agency of β-amylase and their molecular structure (Haworth, Heath & Peat 1942) can be represented by a chain of more than 100 α-glucopyranose members linked as in maltose at the 1 : 4 positions, a portion of such a chain being portrayed in figure 4. It is possible that the chain exists in the form of a helical structure of the type modified by Hanes (1937) (figure 5) from my own studies of the geometric arrangements of six glucopyranose residues which close to form a ring. (This closed ring formula probably represents one of the Schardinger dextrins.) Such a helical structure is also envisaged for the starch-iodine complex (Rundle & French 1943), where each iodine molecule occupies the interior of one coil made up of six glucopyranose members in the continuing convolution of the helix. Under certain experimental conditions the spiral may uncoil to give an extended chain structure.

Sharply distinguished from amylose is the major component of starch: amylopectin. As an approximation to the structure of amylopectin we have evolved the concept of a branched chain or laminated formula. In this the short basal chain, which is the repeating unit, consists of twenty glucose members linked together as in figure 4, these chains being joined in the manner shown diagrammatically in figure 6, where the arrow-heads represent the polymeric linkages. The nature of these linkages is not yet fully understood, but there is reasonable evidence to show that primary valencies are involved and that these are of the 1:6-glucosidic type.

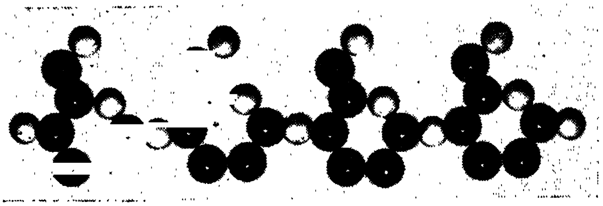


FIGURE 4

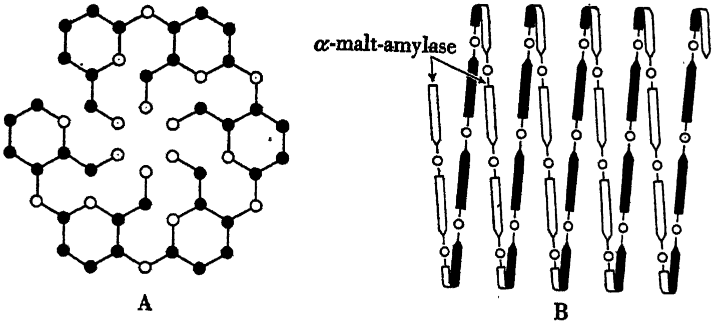


FIGURE 5. A, hexagonal arrangement of six α-linked glucopyranose units (Sachse conformation) modified after Haworth (1939). B, hypothetical spiral model of α-linked chain of 30 glucose units, showing possible mode of combination with α-malt-amylase.

(Taken from a publication by C. S. Hanes in *The New Phytologist*, 1937, **36**, 231.)

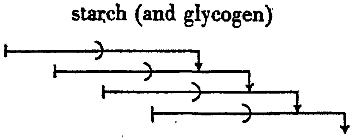


FIGURE 6

The laminated structure serves well to explain certain facts concerning the amylolytic degradation of amylopectin (Haworth, Kitchen & Peat 1943) (figure 7). Under the action of β-amylase this component of starch yields maltose and a limit dextrin, 'dextrin-A', which is not further attacked by β-amylase until it has been 'sensitized' by brief contact with salivary amylase. Thereafter the action of β-amylase continues until a second resting stage is reached, giving 'dextrin-B'.

This dextrin-B is not attacked by β-amylase even after sensitization with saliva, but it is attacked by longer contact with salivary amylase, giving maltose and a 'dextrin-C'. The latter is slowly hydrolysed by pancreatic amylase to 'dextrin-D'. The lengths of the branching chains in dextrins B, C, D are 7–8, 5–6, and 4–5 glucose members respectively. It seems clear that the impediment to β-amylase action is represented by the polymeric linkages mentioned above. Further, the polymeric linkages are broken by a yeast glucosidase (Meyer 1940) and by an enzymic constituent of saliva and of malt α-amylase. It has been postulated that by some as yet unrecognized agency the liberated chains immediately recombine with the formation of new polymeric linkages having a different orientation in the respective chains (figure 8). This may conceivably be effected by a 'Q'-enzyme of the type mentioned below.

Glycogen has many relationships to starch. It gives rise to maltose by enzymic action and its structure is represented by a chain of twelve α-glucopyranose members united by lateral linkages to form a laminated structure similar to that of amylopectin. Thus in figure 6 (above) the curved lines indicate the limits of shorter

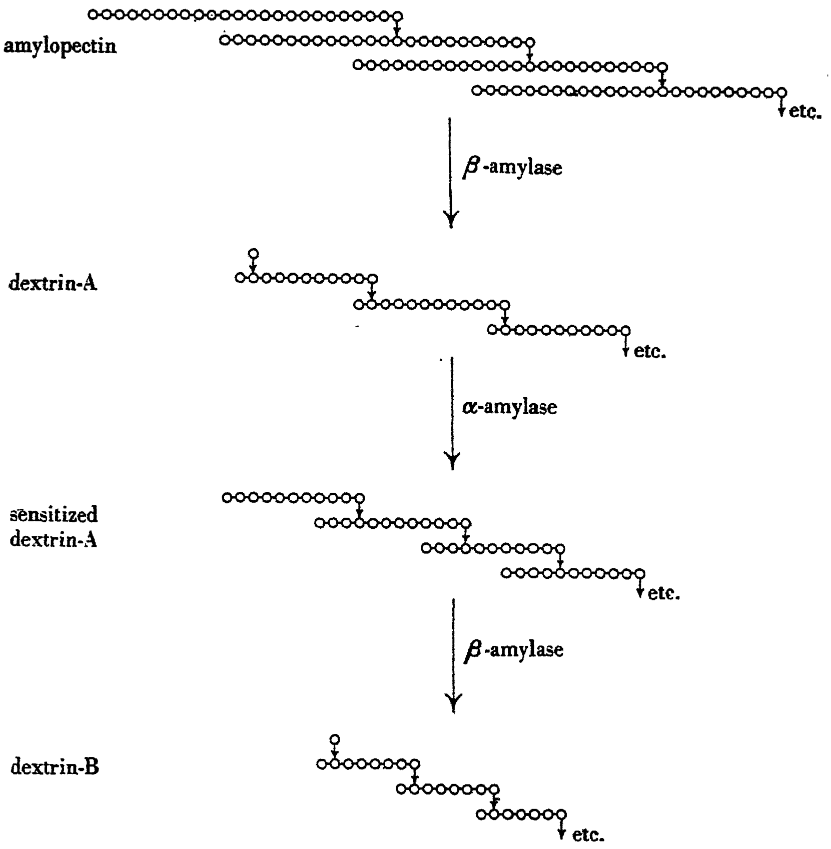


Figure 7. Degradation of starch by amylases.

glycogen chains. Although composed of shorter chain units, glycogen has a molecular weight of from one to two millions, higher than that of starch. It gives a reddish or orange coloration with iodine. Neither starch nor glycogen has the complete symmetry of structure exhibited by cellulose.

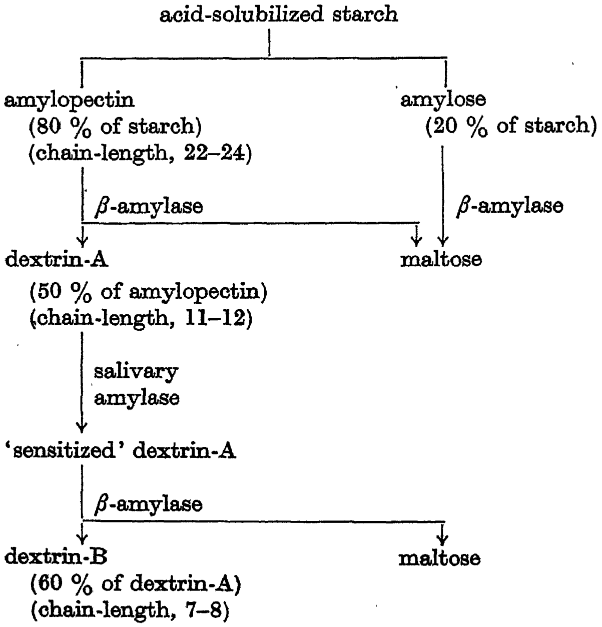


FIGURE 8

The mechanism of the synthesis of a starch polysaccharide by enzymes is of absorbing interest. It is emphasized that these synthetic products prepared *in vitro* reveal that they are very similar to amylose. There are, however, indications that other polysaccharides closely related to glycogen can be synthesized by a liver extract which contains phosphorylases. Latterly, the endeavour in my laboratory to achieve the synthesis of both components of natural starch has given reasonable proof of the synthesis of amylopectin by another agent described as the 'Q'-enzyme to be found in potato. The product obtained by Bourne, Haworth & Peat (1944) has been studied in some detail and has many properties in common with amylopectin (figure 9). Its repeating chain-unit is composed of twenty α-glucopyranose members, and the polysaccharide behaves towards β-amylase exactly as does amylopectin.

Cori and his colleagues, who have devoted much care to the study of animal phosphorylases, have obtained muscle phosphorylase in crystalline form. It is of singular interest that this enzyme requires the nucleotide adenylic acid as a co-enzyme, and the effect on the substrate, glucose-1-phosphate, is the production of an amylose. Liver phosphorylase appears to synthesize a product more closely resembling glycogen.

The importance of hexose phosphates in carbohydrate metabolism has long been known, and their significance as substrates for polysaccharide formation is apparent. Other sugar phosphates, e.g. ribose phosphates, present in nucleic acids and coenzymes, are of high biological significance, and different substituent groups, such as sulphate, may also play important rôles in biological synthesis (see figure 10). Indeed, Jones & Peat have made an interesting speculation on the part played by galactose-1-sulphate in the formation of the polygalactose, agar-agar. The structure of this constituent of a seaweed, of value in bacteriology and industry, has been worked out by Percival and by Peat and their co-workers (Percival and co-workers 1937, 1939; Jones & Peat 1942).

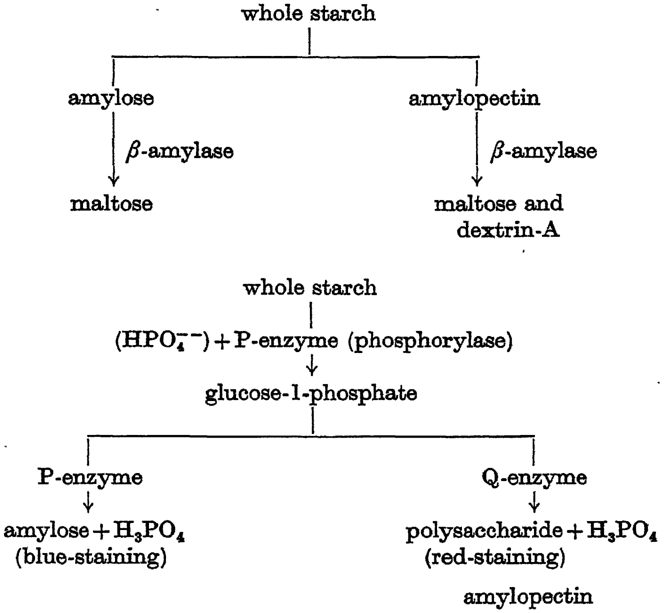
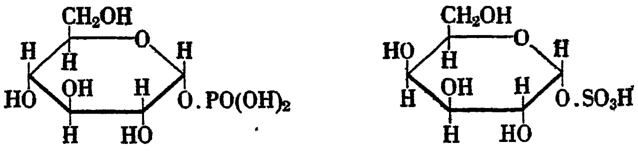
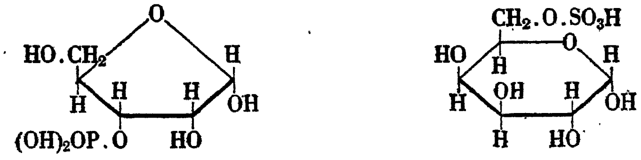


FIGURE 9



Glucose-1-phosphate Galactose-1-sulphate



Ribofuranose-3-phosphate Galactose-6-sulphate

FIGURE 10

polysaccharide	$[\alpha]_D$	equivalent	% N	products of hydrolysis
type I	+294	590	5·2	acetic acid. Amino sugar. Galacturonic acid
II	+ 54	980	0·4	glucose. Aldobionic acid
III	− 33	340	0·3	glucose. Aldobionic acid
IV	+ 30	1550	5·5	acetic acid. Amino sugar

FIGURE 11. Specific polysaccharides of *Pneumococcus*.

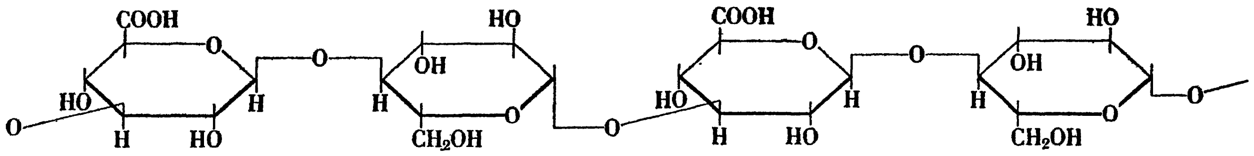


FIGURE 12. Specific polysaccharide, pneumococcus III (Goebel).

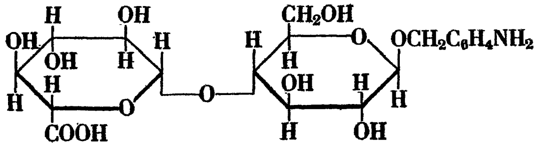


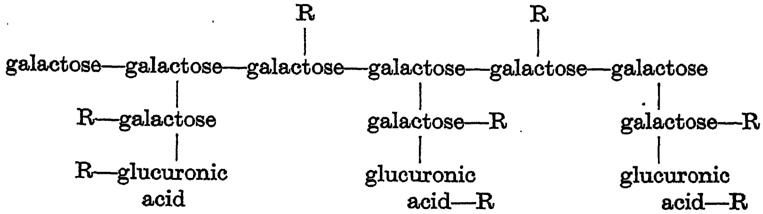
FIGURE 13. *p*-Aminobenzyl *β*-cellobiuronide.

Almost every type of micro-organism possesses enzyme systems which can synthesize polysaccharides, although little is known of these enzymes or of the substrates they utilize. The polysaccharides may act as a cell food reserve, they may form cell-building material, or they may have the function of a protective cell capsule where the filmy outer layer consists of the capsular polysaccharide. When isolated by cautious methods most bacterial polysaccharides possess highly specific physiological properties; they contain hexuronic acids and amino-sugars as constituent units. The nature of the bacterial capsular substance is best exemplified in the *Pneumococcus* group. An extended study initiated by the discoveries of Avery and Heidelberger at the Rockefeller Institute, and occupying more than two decades, has resulted in the differentiation of about forty different types of *Pneumococcus*. The distinction between each type is measured by serological methods and is due to differences in the chemical composition of the capsular polysaccharide. The properties of a few of these polysaccharides are illustrated in figure 11. Of these the type III polysaccharide has been subjected to thorough structural investigation by W. F. Goebel and his colleagues, who have represented the constitution as that of a condensed system of cellobiuronic acid units mutually linked through C_1 of one glucose member with C_3 of the adjoining glucuronic acid member of another unit (figure 12). Goebel (1936) was able to isolate cellobiuronic acid from the type III polysaccharide, and he showed the capacity of this aldobionic acid to determine serological specificity. Preparing *p*-aminobenzyl *β*-cellobiuronide (figure 13), he coupled this by means of the diazo-reaction with normal horse-serum globulin. Using the complex as an antigen (Goebel 1940) he showed that it could confer active and passive immunity on mice not only against type III pneumococcal infection but against types I and II infections as well.

Since degraded products prepared from gum arabic give cross-reactions with pneumococcal antisera (Heidelberger, Avery & Goebel 1929), the aldobionic acid from this plant gum was prepared some years ago in my laboratory and its structure was determined (Challinor, Haworth & Hirst 1931). It is formulated as glucuronic acid-1:6-galactopyranose and this constitution has been confirmed by synthesis (Hotchkiss & Goebel 1936).

The hexuronic acid constituent of aldobionic acid consists simply of a hexose in which the primary alcohol group at C_6 has been oxidized to a carboxyl, those commonly occurring being glucuronic acid (from plant and bacterial gums), galacturonic acid (from pectin) and mannuronic acid (from alginic acid). The complexity of the problems confronting the chemist who is investigating the constitution of the gum polysaccharides is well illustrated by the constitution ultimately assigned to the repeating unit in gum arabic (figure 14) by my colleague, Dr F. Smith, and by the researches of Hirst on other plant gums.

The relationship between serological specificity and chemical structure in the carbohydrate group, as shown by the reaction of pneumococcal antisera with plant gums and aldobionic acid (Marrack 1941), has received further support from the work of Heidelberger & Hobby (1942). These investigators showed that any oxycellulose could react in high dilution with *Pneumococcus* antisera types III and VIII. Interesting also in this connexion is the work on the capsular polysaccharide from *Rhizobium radicicola* (Clover strain). This non-pathogenic nitrogen-fixing organism lives in symbiosis with Leguminosae. Its mucilaginous capsular material, isolated as a thread-like substance, probably has the function of a defence mechanism against soil protozoa, and can readily be synthesized in artificial culture. It reacts in high dilution with type III *Pneumococcus* antiserum. Its constitutional components are shown in figure 15 and its relationship to the type III polysaccharide is apparent (Stacey & Schlüchterer 1945).



R = arabofuranose, or rhamnopyranose, or galactosido-arabofuranose

FIGURE 14. Gum arabic (repeating unit).

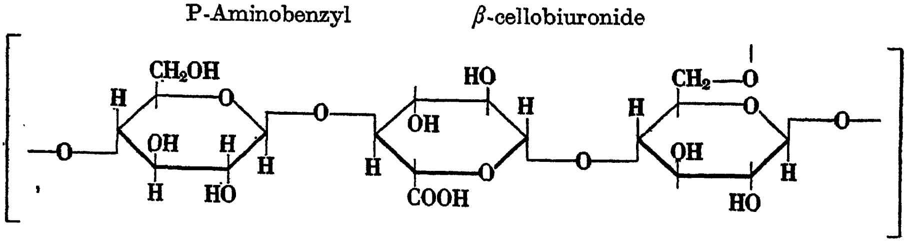


FIGURE 15. Rhizobium polysaccharide.

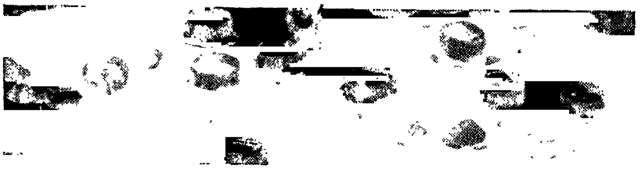


FIGURE 16

Another kind of capsular material is that produced from sucrose by smooth strains of *Betacocci* (*Leuconostoc* species) and by strains of *Betabacteria*, colonies of which on agar plates have the unusual appearance shown in figure 16 (see Mayer 1938). The capsular material encrusting the colonies is composed of a polyglucose which is described as a dextran and it possesses a high dextro-rotation.

The structure which has been determined for the dextrans as a class is that of an extended chain in which the α-glucose members are linked through the 1:6-positions (figure 17) (Peat, Schlüchterer & Stacey 1939). The dextran from *L. dextranicum* has a chain length of *c.* 200 glucose units (Stacey & Daker 1939), that from *Betabact. vermiformé* has a shorter chain length, while that from *L. mesenteroides* prepared by Hibbert has a branched chain structure.

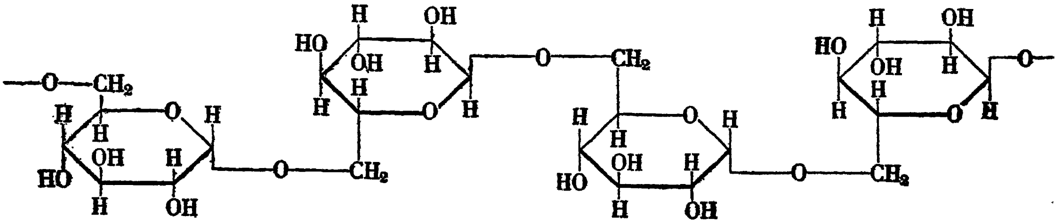


FIGURE 17

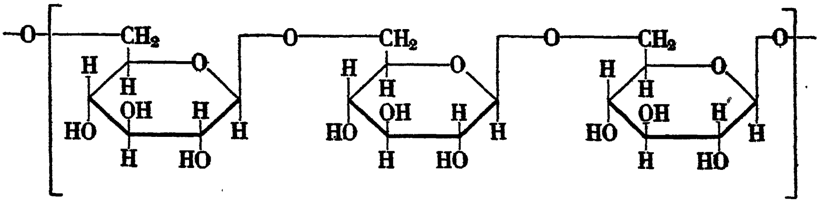


FIGURE 18

A dextran structure has been assigned to luteose, a polyglucose produced from glucose by the mould *Penicillium luteum* (*Zukal*) (Anderson, Haworth, Raistrick & Stacey 1939). In this extended chain formula (figure 18) gentiobiose (glucose-1:6-β-glucose) is the repeating unit. This disaccharide has recently been synthesized in high yield by my colleagues Stacey and Smith, who employed a method analogous to that outlined above for cellobiose. Dextrans have been synthesized directly from sucrose by means of cell-free enzymes (Hehre & Sugg 1942; Stacey 1942) but the mode of synthesis is not yet understood. Only the glucose portion of the sucrose molecule is utilized and it is not clear if phosphorylation processes are involved.

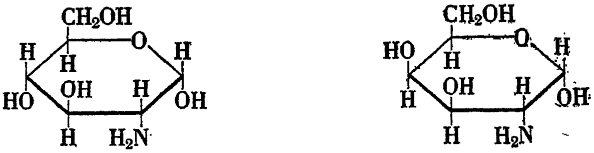
Amino-sugars occur widely in nature as constituents of biologically important polysaccharides. A typical example is the chitin of crustacea and insects, and here the polysaccharide functions as a building material of great strength and stability.

Some bacterial polysaccharides, although protein-free, have a nitrogen content which is often appreciable and is due entirely to the presence of amino-sugar constituents, usually glucosamine (figure 19). Only two naturally occurring amino-sugars are known; these are chitosamine or glucosamine (2-amino-glucose), and chondrosamine (2-amino-galactose) (figure 20). The constitution of glucosamine was determined by chemical methods (Haworth, Lake & Peat 1939) and the configuration of the amino-group on C_2 was confirmed independently by my colleagues

Cox & Jeffrey (1939) who, using the X-ray method, were able to determine accurately the atomic positions in glucosamine hydrobromide without reference to previous stereochemical assumptions. The results proved that glucosamine is a glucose and not a mannose derivative. The decision that chondrosamine is a galactose and not a talose derivative has just been achieved in that it is now shown by

polysaccharide from	% nitrogen
type I *Pneumococcus*	5·5
type IV *Pneumococcus*	6·0
Pneumococcus somatic (group)	5·5
Bact. dysenteriae (Shiga)	2–3
Bact. typhosum	2–3

FIGURE 19. Nitrogen content of polysaccharides of bacterial origin.



2- Amino-glucose
(glucosamine or chitoamine)

2-Amino-galactose
(galactosamine or chondrosamine)

FIGURE 20

A. Mucopolysaccharides
  I. Containing uronic acid
    (*a*) Sulphate-free
      (1) Vitreous humour, umbilical cord, synovial fluid, group A haemolytic *Streptococcus*
      (2) Type I *Pneumococcus*
    (*b*) Sulphate-containing
      (1) Chondroitinsulphuric acid
      (2) Mucoitinsulphuric acid (from gastric 'mucin' and cornea)
      (3) Heparin
  II. Neutral mucopolysaccharides of known composition
    (*a*) Chitin
    (*b*) Gastric polysaccharide
    (*c*) Bacterial polysaccharides (Shiga Kruse (?), type IV *Pneumococcus* (?))
B. Glycoproteins, containing neutral mucopolysaccharides of unknown composition
    (*a*) Ovomucoid-α (formerly called ovomucoid)
    (*b*) Ovomucoid-β (formerly called ovomucin)
    (*c*) Serum mucoid, serum glycoid
    (*d*) Globulins (egg white, serum, thyroglobulin)
    (*e*) Pregnancy urine hormone
C. Nitrogen-free polysaccharides
    Pneumococcus types II and III
    From non-pathogenic bacteria
    From the lower fungi

FIGURE 21

Stacey and his co-workers to be 2-amino-galactopyranose. It gives a series of derivatives analogous to those in the glucosamine series (Stacey 1944). Carbohydrates containing both amino-sugars and hexuronic acids, with other substituents, are often found naturally in firm combination with proteins. These complexes have high physiological significance and constitute the muco-polysaccharides and muco-proteins. A classification of them is shown in figure 21, which has since been modified by Stacey (1943) from the earlier classification by Karl Meyer (1938).

Their physiological function is dependent upon the combination of carbohydrate and protein. Work on this subject by distinguished pioneers (Boivin & Mesrobeanu 1934; Raistrick & Topley 1934) has been continued and advanced, and from *Bact. typhosum* it is possible by means of guanidine salts and other agents to separate the essential immunizing somatic antigen of the cell (Haworth, Jenner, Stacey & Wilson 1938). As shown in figure 22 hydrolytic breakdown of the complex separates a toxic from a specific component. The latter compound is a polysaccharide composed of glucosamine, mannose and galactose units (Stacey 1939) but their precise mode of linking is not yet established. Similar specific poly-

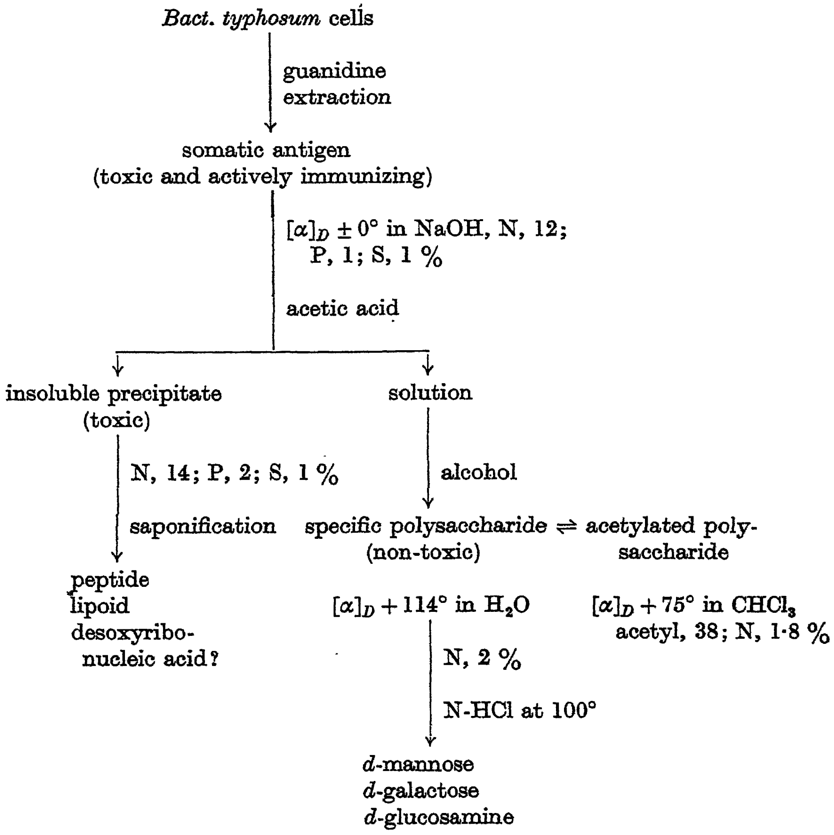

Bact. typhosum cells
guanidine extraction
somatic antigen (toxic and actively immunizing)
$[\alpha]_D$ ± 0° in NaOH, N, 12; P, 1; S, 1 %
acetic acid
insoluble precipitate (toxic)
solution
N, 14; P, 2; S, 1 %
alcohol
saponification
specific polysaccharide (non-toxic)
$\rightleftharpoons$
acetylated polysaccharide
peptide
lipoid
desoxyribonucleic acid?
$[\alpha]_D + 114°$ in H_2O
$[\alpha]_D + 75°$ in $CHCl_3$
acetyl, 38; N, 1·8 %
N, 2 %
N-HCl at 100°
d-mannose
d-galactose
d-glucosamine

FIGURE 22

saccharides exist in other organisms of the Salmonella group (Topley and co-workers 1937; Morgan 1937). The carbohydrate residues constituting the prosthetic groups (the function of which is unknown) in proteins have properties similar to the somatic bacterial polysaccharides, and they all appear to contain glucosamine residues. Ovomucoid from egg protein contains as much as 70 % of this constituent, and the structure of this typical muco-protein has been studied in some detail (Stacey & Woolley 1940, 1942). All the glucosamine units are acetylated at the nitrogen atom and are present as terminal residues giving the unique, highly branched chain structure shown in figure 23.

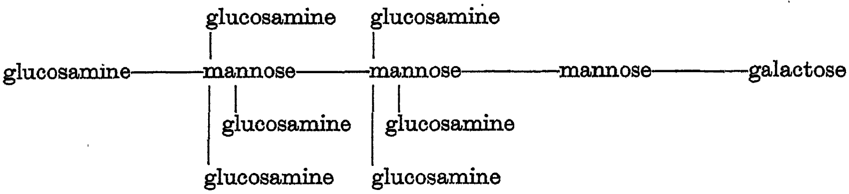


FIGURE 23. Polysaccharide of ovomucoid

It will be seen from the table in figure 21 that some muco-polysaccharides contain sulphated sugar residues in addition to uronic acids and amino-sugars. Some interesting groupings are found in these compounds, the individual units of which may be regarded as types of amino-acid. From hog gastric mucin a neutral polysaccharide having the properties of the specific Blood Group A factor has been shown to contain sugar components identified, in the form of their methylated derivatives, as *l*-fucose, glucosamine, mannose and galactose. From cartilaginous tissue chondroitin sulphate has been prepared (Bray, Gregory & Stacey 1944). This complex carbohydrate, the only known source of chondrosamine, functions as one of the body plastics and it has been shown by Karl Meyer (1938) that it can form salts with proteins such as edestin to reconstitute fibres closely resembling those found in connective tissue. A methylated degraded chondroitin has been examined in my laboratory and one is able to picture the repeating unit in chondroitin as being of the type shown in figure 24, while the fibrous tissue itself is probably constituted along the lines shown in figure 25.

Heparin, the blood anticoagulant of the body, contains glucosamine, glucuronic acid, and sulphate residues. On the basis of analytical studies of a crystalline barium heparinate Wolfrom and his co-workers (1943) have postulated structures of the kind shown in figure 26.

The firm resistance of heparin to acid hydrolysis, and the apparent absence of acetyl residues have led Wolfrom to suggest that some of the glycosidic linkages engage the nitrogen atoms. In all other instances the naturally occurring amino-hexoses have acetyl residues attached to the amino-groups.

Polymers containing pentose sugars occur chiefly in the hemicellulose group, but one of the most important members of the pentose series, *d*-ribose, occurs in

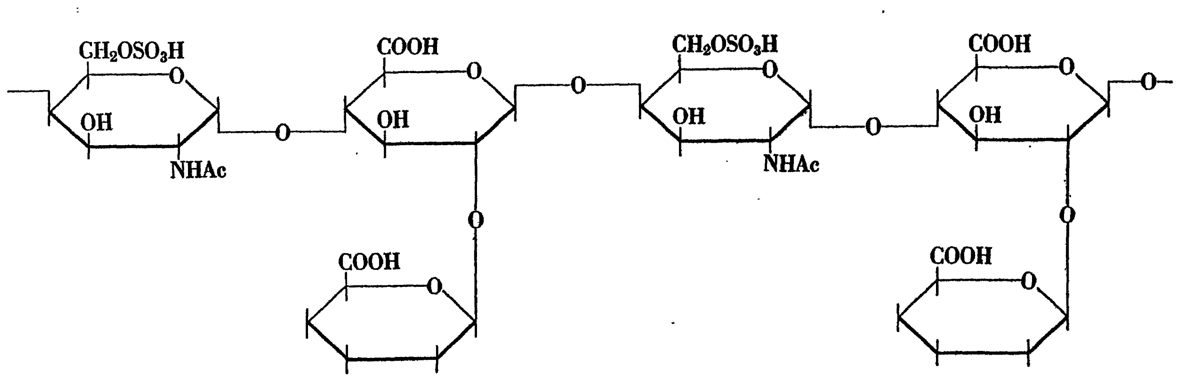


Figure 24. Scheme for chondroitin sulphuric acid.

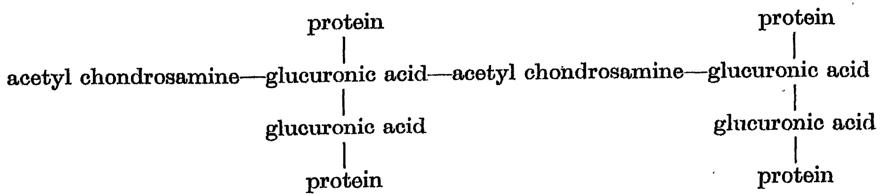

protein
acetyl chondrosamine—glucuronic acid—acetyl chondrosamine—glucuronic acid
glucuronic acid
protein
protein
glucuronic acid
protein

Figure 25. Chondroitin from cartilage.

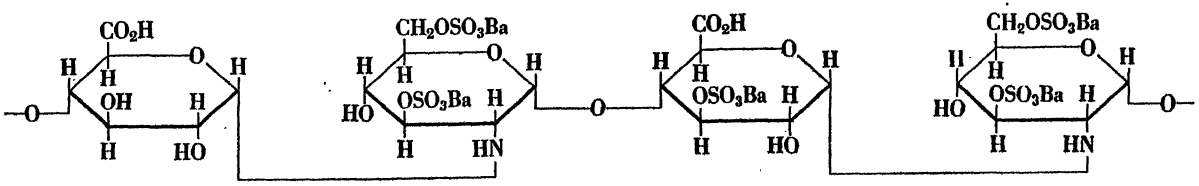


Figure 26

the furanose form in the highly complex nucleic acids. Two types of these, the ribo- and the desoxyribo-nucleic acid, are recognized as being of high significance in cell generation and metabolism. The tentative structures suggested by Levene must merely be regarded as first approximations. Recent investigations on yeast and on some gas gangrene bacilli (Henry & Stacey 1943) have revealed that ribonucleic acid forms an essential constituent of the gram-positive complex. Thus when gram-positive *Cl. welchii* cells (figure 27) are extracted by sodium cholate they assume the appearance of gram-negative cells (figure 28), but the gram-positive character can readily be restored (figure 29) by depositing the magnesium salt of ribo-nucleic acid on the stripped cells. Certain organisms such as *Streptococcus* (shown growing in the smooth phase, figure 30) in the living stage can be made to assume a gram-negative character by starving the organisms of mag-

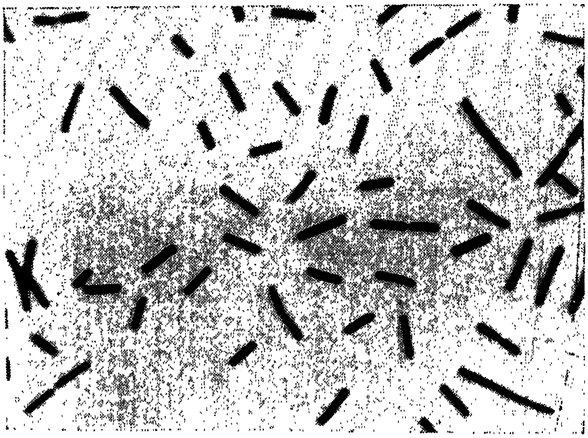


Figure 27

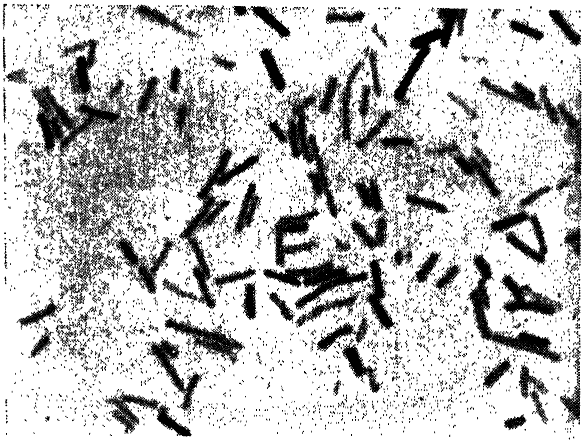


Figure 28

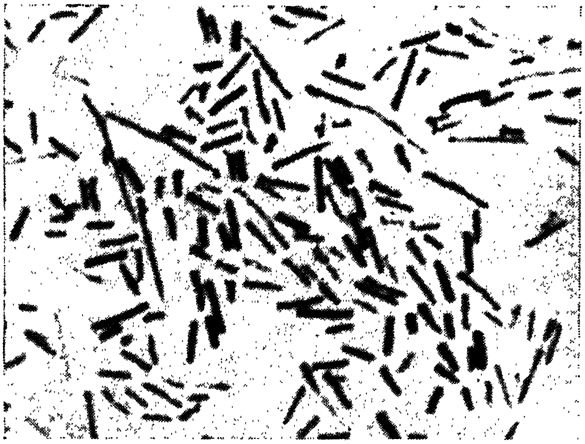


FIGURE 29

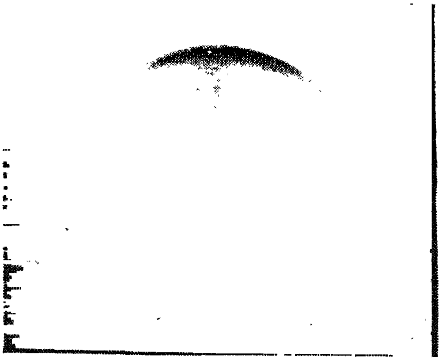


FIGURE 30

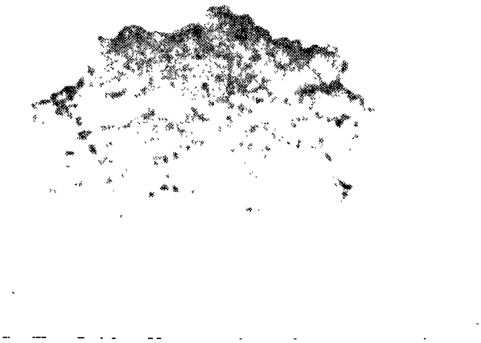


FIGURE 31

nesium. At the same time the colonies change to a rough phase (figure 31). This rough phase indicates not only a lack of ribo-nucleic acid but also a failure of the organism to synthesize its capsular polysaccharide. Both constituents can be restored by subculturing the rough organism on the magnesium-rich medium.

In the case of *Streptococcus salivarius* and related streptococci such as *S. viridans*, the capsular material consists mainly of the polyfructofuranose known as levan. Levans are synthesized only from sucrose and raffinose by a wide variety of micro-organisms, such as the plant pathogens or their extracellular enzymes. The individual polysaccharides differ markedly in their physical characteristics but the studies of Stacey have shown that they all possess the same repeating unit, an extended chain of fructofuranose residues linked through the 2:6-positions (figure 32). They thus appear to be similar to the plant levans, and differ therefore from inulin which possesses a chain of fructofuranoses linked through positions 1 and 2 (figure 33).

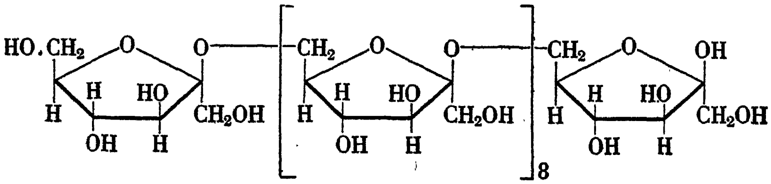


FIGURE 32

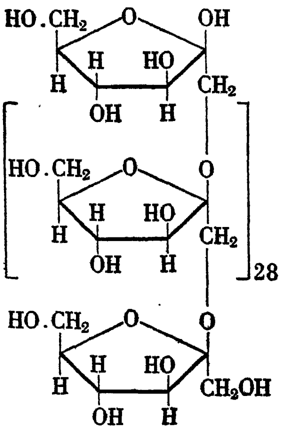


FIGURE 33

The purpose of my lecture has been to demonstrate the wide significance of polysaccharides in the multitudinous variety of problems confronting the chemist and the biologist, and to direct attention to the rich fields which await cultivation by future endeavour.

REFERENCES

Anderson, Haworth, Raistrick & Stacey 1939 *Biochem. J.* **33**, 272.
Bourne, Haworth & Peat 1944 *Nature,* **154**, 236.
Boivin & Mesrobeanu 1934 *C.R. Soc. Biol., Paris,* **115**, 304, 309.
Bray, Gregory & Stacey 1944 *Biochem. J.* **38**, 142.
Challinor, Haworth & Hirst 1931 *J. Chem. Soc.* p. 258.
Cori & Cori 1939 *J. biol. Chem.* **131**, 397.
Cox & Jeffrey 1939 *Nature,* **143**, 894.
Goebel 1936 *J. Exp. Med.* **64**, 29.
Goebel 1940 *J. Exp. Med.* **72**, 33.
Hanes 1937 *New Phytol.* **36**, 101.
Hanes 1940 *Proc. Roy. Soc.* B, **129**, 174.
Haskins, Hann & Hudson 1942 *J. Amer. Chem. Soc.* **64**, 1289.
Hassid & McCready 1943 *J. Amer. Chem. Soc.* **65**, 1157.
Haworth 1939 *Chem. and Ind.* **58**, 917.
Haworth, Heath & Peat 1942 *J. Chem. Soc.* p. 55.
Haworth, Jenner, Stacey & Wilson 1938 Unpublished results.
Haworth, Kitchen & Peat 1943 *J. Chem. Soc.* p. 619.
Haworth, Lake & Peat 1939 *J. Chem. Soc.* p. 271.
Hehre & Sugg 1942 *J. Exp. Med.* **75**, 339.
Heidelberger, Avery & Goebel 1929 *J. Exp. Med.* **49**, 847.
Heidelberger & Hobby 1942 *Proc. Nat. Acad. Sci., Wash.,* **28**, 516.
Henry & Stacey 1943 *Nature,* **151**, 671.
Hibbert 1931 *Can. J. Res.* **5**, 580.
Hotchkiss & Goebel 1936 *J. Biol. Chem.* **115**, 285.
Jones & Peat 1942 *J. Chem. Soc.* p. 225.
Kerr 1942 *Paper Trade J.* **115**, 30.
Marrack 1941 *Chemistry of antigens and antibodies.* London: Medical Research Council.
Mayer 1938 *Das 'Tibi' Konsortium.* Delft.
Meyer 1938 *Cold Spr. Harb. Symp.* **6**, 91.
Meyer 1940 *Helv. Chim. Acta,* **23**, 875.
Morgan 1937 *Biochem. J.* **31**, 2003.
Peat, Schlüchterer & Stacey 1939 *J. Chem. Soc.* p. 581.
Percival & co-workers 1937 *J. Chem. Soc.* p. 1615.
Percival & co-workers 1939 *J. Chem. Soc.* p. 1844.
Raistrick & Topley 1934 *Brit. J. exp. Path.* **15**, 113.
Rundle & French 1943 *J. Amer. Chem. Soc.* **65**, 1707.
Schoch 1942 *J. Amer. Chem. Soc.* **64**, 2951.
Stacey 1939 (in the Press).
Stacey 1942 *Nature,* **149**, 639.
Stacey 1943 *Chem. and Ind.* **62**, 110.
Stacey 1944 *J. Chem. Soc.* p. 272.
Stacey & Daker 1939 *J. Chem. Soc.* p. 584.
Stacey & Schlüchterer 1945 *J. Chem. Soc.* p. 776.
Stacey, Smith & Gilbert 1944 *Chem. and Ind.* p. 423.
Stacey & Woolley 1940 *J. Chem. Soc.* p. 184.
Stacey & Woolley 1942 *J. Chem. Soc.* p. 550.
Topley and co-workers 1937 *Lancet,* p. 252.
Wolfrom and co-workers 1943 *J. Amer. Chem. Soc.* **65**, 2077.

Turbulence and diffusion in the lower atmosphere

By R. Frost, B.A.

(*Communicated by E. Gold, F.R.S.—Received* 16 *December* 1944)

The only existing theory of atmospheric turbulence which is capable of giving a quantitative approach to the complex problems of diffusion in the lower atmosphere is the classical theory in which it is generally assumed that the effect of eddies in the atmosphere is completely analogous to that of molecules in a gas apart from a difference of scale. This assumption, which later evidence has shown to be incorrect, is not essential to the theory, and in the present paper is replaced by the assumption that the mixing length of an eddy increases with both height above and nature of the earth's surface.

With this assumption a self-consistent treatment of diffusion is developed which is able to account quantitatively for such meteorological phenomena as the distribution of water vapour over land and sea (including evaporation from the oceans) and the diffusion of smoke near the ground. The treatment is mainly confined to diffusion in an adiabatic atmosphere.

1. Introduction

The classical treatment of diffusion in the atmosphere, which is largely due to Taylor and Richardson in England and Schmidt in Austria, showed that it was possible to give a rational explanation to many of the complex problems of diffusion in the atmosphere by replacing the molecular coefficient of diffusion of the order of 10^{-1} cm.^{2}/sec. by a virtual coefficient of diffusion of the order of 10^3–10^5 cm.^{2}/sec. The first and natural assumption that the virtual coefficient was a constant was unsatisfactory and measurements of K, the virtual coefficient of diffusion, showed that in the lowest few hundred metres above the earth's surface it increased rapidly with height, being of the order of 10^{-1} cm.^{2}/sec. a few cm. above the ground and increasing to 10^4–10^5 cm.^{2}/sec. at 300–500 m. above the ground.

In the first section of the present paper a semi-empirical expression is obtained for K, and in subsequent sections with the aid of this expression solutions of the equations of eddy diffusion and the equations of motion with given boundary conditions are obtained which are self-consistent and which are in good agreement with observations. In the absence of a satisfactory dynamical theory of turbulence* it is probable that this method of attack gives the most profitable approach to the complex problems of diffusion near the earth's surface.

2. The derivation of the power laws

Let V be the mean velocity of the air at a height z above the surface of the earth and ρ the density of the air.

According to Prandtl's development of Taylor's theory the coefficient of eddy diffusion may be written

$$K = l^2 \frac{\partial V}{\partial z}, \tag{2·1}$$

---

* For a good general exposition of the present position of the theory of turbulence see Goldstein (1938), and for its application in meteorology, Brunt (1939).

and τ, the eddy shearing stress, by

$$\tau = \rho l^2 \left(\frac{\partial V}{\partial z}\right)^2, \tag{2·11}$$

where l is a length of somewhat indefinite nature which he defines as the mean vertical distance traversed by an eddy before it mixes with its environment.

Owing to the lack of knowledge of the dynamical processes involved in the motion of an individual eddy from one layer to another it is not possible to derive a theoretical expression for l in terms of measurable quantities. Observations of pipe flow, however, suggest that l is independent of the velocity and depends only upon the distance from the nature of the boundary surface.

Rossby & Montgomery in their work assume that over a rough surface and in an adiabatic atmosphere l can be represented by

$$l = a(z+z_0), \tag{2·2}$$

where a is a non-dimensional constant and z_0 is a length proportional to the mean height of the surface irregularities. This particular expression for l leads to difficulties when dealing with the differential equations involved in the problems of turbulent transfer near the earth's surface, and furthermore, observations of flow in tubes by G. I. Taylor, Goldstein and others, show that the so-called constant a decreases with increasing distance from the boundary.

In the present paper, as a working hypothesis, it will be assumed that over a rough surface

$$l = z^{1-m} z_0^m, \tag{2·3}$$

where m is a non-dimensional constant and z_0 is a length characteristic of the degree of roughness of the surface. For l to increase with both height above and roughness of the surface it follows that m must lie between 0 and 1.

Equation (2·3) gives for K and τ

$$K = z^{2-2m} z_0^{2m} \frac{\partial V}{\partial z}, \tag{2·4}$$

$$\tau = \rho z^{2-2m} z_0^{2m} \left(\frac{\partial V}{\partial z}\right)^2. \tag{2·41}$$

Now Ertel (1933) has shown that in the lowest 100 ft. or so of the atmosphere τ/ρ is very nearly a constant, and hence, taking the square root of (2·41) and integrating, one obtains if V and z vanish together,

$$V = \frac{1}{m}\sqrt{\frac{\tau}{\rho}}\left(\frac{z}{z_0}\right)^m \tag{2·5}$$

or

$$V = V_h \left(\frac{z}{h}\right)^m, \tag{2·51}$$

where V_h is the mean velocity measured at a standard height h.

Observations by many investigators (summarized by Carruthers 1943) have shown that the variation of wind with height near the surface of the earth can be expressed by a power law such as (2·51), where m can have any value between 0 and 1 and appears to be mainly, if not entirely, a function of the stability. Thus when the lapse rate is very large and negative m tends to 0 and when there is a large inversion m tends to the value 1. Under conditions of thermal equilibrium in wind tunnels $m = \frac{1}{7}$ approximately, and the same is true in the atmosphere. With strong winds M. A. Giblett's observations at Cardington indicate that $m = \frac{1}{7}$ apparently irrespective of the lapse rate.

Thus, write

$$K = mz^{1-m} z_0^{2m} V_h h^{-m}, \tag{2·6}$$

$$\tau = \rho m^2 z_0^{2m} V_h^2 h^{-2m}, \tag{2·61}$$

and for purposes of computation take $m = \frac{1}{7}$. For comparison, observations made only when the winds are moderate or strong will be used.

Rossby, from wind measurements by Wust (1920) at different heights above the sea, found with moderate and strong winds, when the sea has the character of a rough surface, that τ, the shearing stress, could be represented by

$$\tau = 2{\cdot}6 \times 10^{-3} \rho V_h^2, \tag{2·7}$$

where V_h was measured at a height of 15 m. Ekman derived a similar result as early as 1905 by an entirely different approach.

If (2·7) and (2·61) be identified it can be seen that

$$2{\cdot}6 \times 10^{-3} \rho V_h^2 = 2{\cdot}5 \times 10^{-3} z_0^{\frac{2}{7}} \rho V_h^2,$$

whence

$$z_0 = 1\,\text{cm.} \tag{2·71}$$

G. I. Taylor, from pilot balloon observations by Dodson over Salisbury Plain, obtained for the shearing stress with moderate and strong winds

$$\tau = 2{\cdot}7 \times 10^{-3} \rho V_h^2, \tag{2·8}$$

where V_h was measured at a height of 30 m. above the ground.

If this be identified with (2·61) a value of 2·6 cm. is obtained for z_0. (2·81)

K in (2·6) was derived as a coefficient of eddy viscosity. Observations by Sverdrup (1936) of wind velocities, temperatures and vapour pressures at various heights over a snow surface, however, show that the coefficients of eddy viscosity, eddy diffusivity and eddy conductivity obey the same power laws, and hence in subsequent sections this expression for K will be used as a general coefficient of diffusion.

3. THE EVAPORATION PROBLEM

It is assumed that the water surface is flush with the ground over which the air stream has been passing, and that the moisture content of the air initially over the land is invariant with height. Let χ be the quantity of water vapour per unit volume of air in excess of that present in the air stream before it reaches the water surface, and let x be measured downwind from the windward edge of the water surface. Neglecting horizontal diffusion, the equation stating that advection and diffusion balance is

$$V\frac{\partial \chi}{\partial x} = \frac{\partial}{\partial z}\left\{K\frac{\partial \chi}{\partial z}\right\}, \tag{3·1}$$

which may be written

$$V_h h^{-m} z^m \frac{\partial \chi}{\partial x} = \frac{\partial}{\partial z}\left\{m V_h h^{-m} z_0^{2m} z^{1-m}\frac{\partial \chi}{\partial z}\right\} \tag{3·11}$$

or

$$\frac{\partial \chi}{\partial x} = a z^{-m}\frac{\partial}{\partial z}\left\{z^{1-m}\frac{\partial \chi}{\partial z}\right\}, \tag{3·12}$$

where $a = m z_0^{2m}$.

It is assumed that evaporation takes place in such a way that $\chi = \chi_0$ a constant, over the water surface, i.e. over the area for which $z = 0$ and x is positive.

Examination of the dimensions of the terms in (3·12) suggests that it should be possible to obtain for that equation a solution in which χ is a function of the variable ξ defined by

$$\xi = \frac{z^{2m+1}}{ax}. \tag{3·2}$$

Denoting differentiation with respect to ξ by dashes, such a solution must satisfy

$$\{(2m+1)(m+1)+\xi\}\chi' + (2m+1)^2\xi\chi'' = 0. \tag{3·3}$$

The first integral of this equation is

$$\chi' = -\chi_0 G \xi^{-\frac{m+1}{2m+1}} \exp\left\{-\frac{\xi}{(2m+1)^2}\right\}, \tag{3·4}$$

where G is a constant.

Now $\xi \to \infty$ both when $x \to 0$ and when $z \to \infty$. In either case χ is to be zero. It follows that all the conditions of the problem are satisfied by the solution

$$\chi = \chi_0 G \int_\xi^\infty \xi^{-\frac{m+1}{2m+1}} \exp\left\{-\frac{\xi}{(2m+1)^2}\right\} d\xi, \tag{3·5}$$

provided that

$$G = \left[(2m+1)^{\frac{2m}{2m+1}}\Gamma\left(\frac{m}{2m+1}\right)\right]^{-1}. \tag{3·51}$$

Writing $\xi = (2m+1)^2\theta$ equation (3·5) becomes

$$\chi = \frac{\chi_0 \int_\theta^\infty \theta^{-\frac{m+1}{2m+1}} e^{-\theta}\, d\theta}{\Gamma\left(\frac{m}{2m+1}\right)}. \tag{3·52}$$

Thus χ is determined by an incomplete Γ function. O. G. Sutton (1934), working with the same assumptions, obtained for χ an expression involving an integral of the modified Bessel function which was not suitable for mathematical computation.*

For $m = \frac{1}{7}$, the ratios χ/χ_0 for varying values of θ are given in table 1, and are shown graphically in figure 1.

TABLE 1

θ	0	10^{-12}	10^{-11}	10^{-10}	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}
χ/χ_0	1·000	0·9510	0·9367	0·9182	0·8944	0·8636	0·8239	0·7724	0·7062	0·6206
θ	10^{-3}	2×10^{-3}	3×10^{-3}	4×10^{-3}	5×10^{-3}	6×10^{-3}	7×10^{-3}	8×10^{-3}	9×10^{-3}	10^{-2}
χ/χ_0	0·5099	0·4707	0·4463	0·4285	0·4143	0·4022	0·3918	0·3828	0·3749	0·3676
θ	10^{-2}	2×10^{-2}	3×10^{-2}	4×10^{-2}	5×10^{-2}	6×10^{-2}	7×10^{-2}	8×10^{-2}	9×10^{-2}	10^{-1}
χ/χ_0	0·3676	0·3176	0·2869	0·2645	0·2467	0·2321	0·2196	0·2087	0·1990	0·1903
θ	10^{-1}	2×10^{-1}	3×10^{-1}	4×10^{-1}	5×10^{-1}	6×10^{-1}	7×10^{-1}	8×10^{-1}	9×10^{-1}	1·0
χ/χ_0	0·1903	0·1337	0·1019	0·0807	0·0654	0·0538	0·0448	0·0376	0·0318	0·0271
θ	1·1	1·2	1·3	1·4	1·5	1·6	1·7	1·8	1·9	2·0
χ/χ_0	0·0231	0·0198	0·0171	0·0147	0·0128	0·0111	0·0096	0·0084	0·0073	0·0064

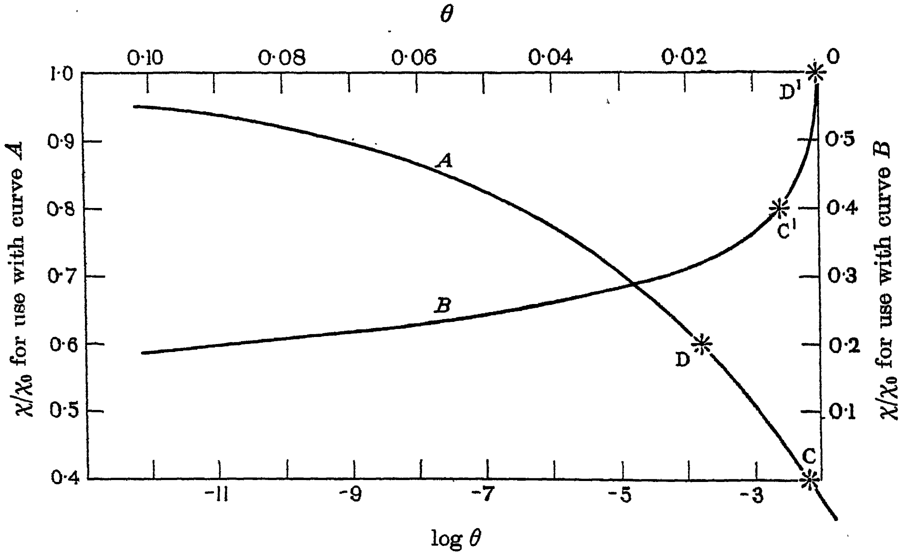


FIGURE 1. Curve A; values of χ/χ_0 are plotted against values of $\log\theta$ for values between $\theta = 10^{-12}$ and $\theta = 10^{-3}$. Curve B; values of χ/χ_0 are plotted against values of θ between $\theta = 10^{-3}$ and $\theta = 10^{-1}$. It will be seen that CD on curve A corresponds to $C'D'$ on curve B.

---

[*] The above solution was communicated to Mr O. G. Sutton early in 1934 and later in the same year to the late Dr F. J. W. Whipple, who showed by a previously unknown identity that Sutton's solution (with two errata corrected) could be reduced to the above form. The identity involved is

$$\int_0^\infty t^{1-p}K_p(t)\exp\left(-\frac{t_2}{4c}\right)dt = 2^{-p}\Gamma(1-p)\,c^{1-p}e^c\int_0^\infty e^{-u}u^{p-1}du.$$

. Since with $m = \frac{1}{7}$ and $z_0 = 1$ cm. and therefore

$$\theta = \frac{z^{\frac{9}{7}} \times 49 \times 7}{81 \times x} = \frac{4{\cdot}23 z^{\frac{9}{7}}}{x},$$

it is a simple matter with the aid of table 1, or figure 1, to find the increase of water vapour at any given height above the sea surface after any specified distance of travel. For example, with initially dry air, the relative humidity at a height of 1·2 m. (4 ft.) above the surface of the sea would be 62 % when $\theta = 10^{-4}$ or when

$$\begin{aligned} x &= 4{\cdot}23 \times (120)^{\frac{9}{7}} \times 10^4 \\ &= 198\,\text{km}. \end{aligned}$$

Similarly, from figure 1 the relative humidity would be 70 % when

$$\begin{aligned} x &= 4{\cdot}23 \times (120)^{\frac{9}{7}} \times 10^4 \times 9 \\ &= 1800\,\text{km}. \end{aligned}$$

N.B. It should be noted that with moderate and strong winds the increase of water vapour depends only upon the length of sea track and is independent of the wind speed.

The rate at which water vapour is carried away by the air stream can readily be derived from (3·4), for

$$\frac{\partial \chi}{\partial z} = \chi' \frac{\partial \xi}{\partial z} = -\frac{(2m+1)\, G\chi_0}{(ax)^{\frac{m}{2m+1}}} z^{m-1} \exp\left\{-\frac{\xi}{(2m+1)^2}\right\}, \tag{3·6}$$

and hence $e(x)$, the rate of evaporation per unit area at a distance x from the land, is given by

$$e(x) = \underset{z\to 0}{\mathrm{Lt}} -K\frac{\partial \chi}{\partial z} = \frac{(2m+1)^{\frac{1}{2m+1}} x^{-\frac{m}{2m+1}} a^{\frac{m+1}{2m+1}} V_h h^{-m} \chi_0}{\Gamma\left(\dfrac{m}{2m+1}\right)}. \tag{3·61}$$

By integrating the right-hand side of (3·61) with respect to x it is found that the rate of evaporation from a surface of unit breadth and of length x_0 downwind is given by

$$\begin{aligned} E &= \frac{(2m+1)^{\frac{2m+2}{2m+1}} x_0^{\frac{m+1}{2m+1}} a^{\frac{m+1}{2m+1}} V_h h^{-m} \chi_0}{(m+1)\, \Gamma\left(\dfrac{m}{2m+1}\right)} \\ &= \frac{(2m+1)^{\frac{2m+2}{2m+1}} (m x_0 z_0^{2m})^{\frac{m+1}{2m+1}} V_h h^{-m} \chi_0}{(m+1)\, \Gamma\left(\dfrac{m}{2m+1}\right)}. \end{aligned} \tag{3·62}$$

Thus when $m = \frac{1}{7}$ and $z_0 = 1$ cm.

$$E = 2{\cdot}85 \times 10^{-2} x_0^{\frac{8}{9}} \chi_0 V_h h^{-\frac{1}{7}}. \tag{3·63}$$

For small values of θ, (3·52) is to a high order of approximation given by

$$\chi = \chi_0 \left[1 - \frac{\theta^{\frac{m}{2m+1}}}{\Gamma\left(1 + \frac{m}{2m+1}\right)} \right], \tag{3·7}$$

and so
$$\chi_0 - \chi = \frac{\chi_0 z^m}{[(2m+1)^2 ax]^{\frac{m}{2m+1}} \Gamma\left(1 + \frac{m}{2m+1}\right)}. \tag{3·71}$$

Eliminating x between (3·61) and (3·71) we have

$$\begin{aligned} e(x) &= \frac{m^2 z_0^{2m} V_h h^{-m} (\chi_0 - \chi)}{z^m} \\ &= \frac{m^2 z_0^{2m} V_h (\chi_0 - \chi_h)}{h^{2m}}, \end{aligned} \tag{3·8}$$

or writing $m = \frac{1}{7}$, $z_0 = 1$ cm. and $h = 15$ m.

$$e(x) = 2 \times 10^{-9} V_h (p_w - p_h), \tag{3·81}$$

where p_w and p_h are the vapour pressures at the surface of the sea and at a height $h = 15$ m. respectively. Thus the empirical formula $E = k_h(p_w - p_h) V_h$, where k_h is a constant, suggested by many writers for use with moderate and strong winds, can be derived theoretically and a numerical value of the constant obtained.

As an illustration of the use of these equations, consider the evaporation from a large belt of water approximately 5000 km. in length, and with a surface temperature of 49° F, into a turbulent atmosphere in which the mean wind speed at a height of 10 m. is 7 m./sec. The air will be assumed to have initially a mean moisture content of 6·3 g./cu.m. These figures represent approximately the mean annual values over the Atlantic Ocean between latitudes 50 and 55° N. According to equation (3·63) the mean rate of evaporation in mm. per day is

$$\begin{gathered} 2{\cdot}85 \times 10^{-2} \times 700 \times (1000)^{-\frac{1}{7}} \times (5 \times 10^8)^{-\frac{1}{9}} \times (909 - 6{\cdot}3) \times 10^{-6} \times 3600 \times 24 \times 10 \\ = 1{\cdot}87 \text{ mm./day}. \end{gathered}$$

The value of μ, the moisture content at a height of 4 ft. above the surface after a sea crossing of 5000 km., is readily obtained, for with $z = 120$ cm., $x_0 = 5 \times 10^8$ cm. and $z_0 = 1$ cm., $\theta = 4{\cdot}1 \times 10^{-6}$, and hence from table 1

$$\frac{\mu - 6{\cdot}3}{9{\cdot}09 - 6{\cdot}3} = \frac{73}{100}.$$

Therefore, $\mu = 8{\cdot}3$ g./cu.m.

Wust (1920) gives the following values of the mean rate of evaporation from the Atlantic Ocean:

latitude 50–60° N: 1·0 mm./day,
latitude 40–50° N: 1·8 mm./day.

Bilham (1938) gives the mean moisture content at Valentia as 8·3 g./cu.m. and at Falmouth as 8·2 g./cu.m. Thus the observed annual values are in very good agreement with the calculated values.

χ could equally well have been taken to represent the increase of potential temperature of the air over that initially in the air stream before it reaches the water surface provided that the potential temperature of the air initially was invariant with height.

More generally, since Θ, the potential temperature of the air (or μ the moisture content of the air) over the land surface, initially satisfies (in the steady state)

$$\frac{\partial}{\partial z}\left\{K\frac{\partial\Theta}{\partial z}\right\} = 0,$$

where $K \propto z^{1-m}$ it follows that

$$\Theta = \Theta_1 + Bz^m,$$

where Θ, is the potential temperature at the ground.

But since this solution also satisfies the differential equation (3·12), it may be combined with (3·52), and thus

$$\Theta = \Theta_1 + Bz^m + (\Theta_0 - \Theta_1)\int_\theta^\infty \theta^{-\frac{m+1}{2m+1}} e^{-\theta}\, d\theta \tag{3·9}$$

is the solution of the differential equation (3·12), which satisfies the more general boundary conditions.

The following examples taken from the Daily Weather Reports serve to illustrate and confirm the above theory. Thus on 26 March 1939:

'The temperatures of the air at Skagen and Blaavands Huk in Denmark were 32 and 31°, respectively, and the dew-point at both stations was 26° F. The temperature of the North Sea from ships' observations was 43° F. The temperature of the air reaching England was 38° F at Spurn and 39° F at Gorleston and the dew-point of the air had risen to 37° F.'

From figure 1, taking the height above sea surface as 120 cm. and the sea track 600 km., the temperature should have been

$$31\tfrac{1}{2} + 0\cdot63(43-32)$$
$$= 38\tfrac{1}{2}.$$

Now a dew-point of 26° F corresponds with a moisture content of 3·8 g. water vapour/cu.m., and a dew-point of 43° F corresponds with 7·3 g./cu.m. Thus the final moisture content should be

$$3\cdot8 + 0\cdot63(7\cdot3 - 3\cdot8)$$
$$= 6\cdot0 \text{ g./cu.m.}$$

which corresponds to a dew-point of 37° F.

On 10 December 1936:

'The air temperature on the English coast was 33° F, whilst the temperature of the sea from ships' observations was 47° F. The temperature of the air on arrival in Norway was 41° F.'

According to the above theory the temperature should have been

$$33^\circ + 0\cdot63(47-33)$$
$$= 42^\circ\ \mathrm{F}.$$

On 15 December 1936:

'The air temperature leaving Scotland was 36° F. The sea temperature from ships' observations was 47° F, and the air temperature on the Norwegian coast on the 16th was 43° F.'

According to the present theory the final temperature should have been

$$36^\circ + 0\cdot63(47-36)$$
$$= 43^\circ\ \mathrm{F}.$$

From the meteorological aspect the corresponding problem of the distribution of water vapour in an air stream which, after flowing over the sea, flows over a dry surface, is of equal importance. This problem, however, is complicated by the fact that in general the roughness parameter of the land is greater than that of the sea, and that K in consequence changes when the air passes from sea to land. An exact solution may be obtained for the particular case in which the roughness parameters of the land and sea are the same, and is given in § 4. It is not easy to obtain an exact solution of the more general problem, but in § 5 an approximate solution of this is given, which is sufficiently accurate for all practical purposes and which is very suitable for computation.

4. AN INFINITE LINE SOURCE

The problem of the distribution of matter in an air stream due to an infinite line source flush with the ground and at right angles to the air flow, which is continuously emitting, is of considerable importance both from a theoretical and practical aspect.

Let x be measured downwind from the line source and z vertically upwards from the ground, and let χ be the density of suspended matter in g./cu.m. A solution is therefore required of the dimensional equation

$$z^m \frac{\partial \chi}{\partial x} = a_1 \frac{\partial}{\partial z}\left\{z^{1-m}\frac{\partial \chi}{\partial z}\right\}, \qquad (4\cdot1)$$

where $a_1 = m z_{10}^{2m}$, where z_{01} is the characteristic roughness of the surface, which satisfies the following boundary conditions:

(1) $\chi \to 0$, when $x \to 0$, $z \neq 0$.

(2) $\chi \to 0$, when $z \to \infty$.

(3) $\int_0^\infty V\chi\,dz = \phi$, i.e. the quantity of matter passing through a plane at right angles to the line source in unit time is a constant and equal to the rate of emission at the source. ϕ is the rate in g./sec./unit length of the line source.

As in § 3, an examination of the dimensions of the equation suggests that χ is a function of θ, where

$$\theta = \frac{z^{2m+1}}{(2m+1)^2 a_1 x}.$$

A simple solution $\chi = f(\theta)$ gives

$$\phi = \int_0^\infty V_h h^{-m} z^m f(\theta)\, d\theta$$

$$= A x^{\frac{1+m}{1+2m}} \int_0^\infty f(\theta)\, d\theta^{\frac{1+m}{1+2m}},$$

which depends upon x. It suggests, however, that it should be possible to find a solution

$$\chi = x^{-\frac{1+m}{1+2m}} f(\theta). \tag{4·2}$$

Denoting differentiation with respect to θ by dashed letters, $f(\theta)$ must satisfy

$$\theta f'' + \left(\theta + \frac{1+m}{1+2m}\right) f' + \frac{1+m}{1+2m} f = 0 \tag{4·3}$$

or

$$\left(\theta \frac{\partial}{\partial\theta} + \frac{1+m}{1+2m}\right)\left(\frac{\partial}{\partial\theta} + 1\right) f = 0,$$

whence

$$f(\theta) = B e^{-\theta} + C e^{-\theta} \int e^{\theta}\, \theta^{-\frac{1+m}{1+2m}}\, d\theta. \tag{4·4}$$

Since, however, $\chi \rightarrow 0$ when $\theta \rightarrow \infty$, it follows that $C = 0$, and hence

$$\chi = B x^{-\frac{m+1}{2m+1}} e^{-\theta} \tag{4·41}$$

clearly satisfies all the conditions of the problem providing

$$\phi = \int_0^\infty B e^{-\theta} x^{-\frac{m+1}{2m+1}} V_h h^{-m}\, d\left(\frac{z^{m+1}}{m+1}\right)$$

$$= B V_h h^{-m} (2m+1)^{\frac{1}{2m+1}} a_1^{\frac{m+1}{2m+1}} \Gamma\left(\frac{m+1}{2m+1}\right). \tag{4·5}$$

Thus

$$\chi = \frac{\phi x^{-\frac{m+1}{2m+1}} \exp\left\{-\dfrac{z^{2m+1}}{(2m+1)^2 a_1 x}\right\}}{V_h h^{-m} (2m+1)^{\frac{1}{2m+1}} a_1^{\frac{m+1}{2m+1}} \Gamma\left(\dfrac{m+1}{2m+1}\right)}. \tag{4·6}$$

Thus the concentration of smoke downwind at the ground should vary as $x^{-\frac{m+1}{2m+1}}$ or with $m = \frac{1}{7}$ as $x^{-0{\cdot}89}$.

From (4·6) it can be seen that the distribution of smoke with height at a distance x downwind is given by

$$\chi(x, z) = \chi(x, 0) \exp\left\{-\frac{z^{2m+1}}{(2m+1)^2 m z_{01}^{2m} x}\right\}, \tag{4·7}$$

or with $m = \frac{1}{7}$ and $z_{01} = 2{\cdot}6$

$$\chi(x,z) = \chi(x,0)\exp\left\{-\frac{3{\cdot}223z^{\frac{9}{7}}}{x}\right\}.$$

The height of a smoke cloud is usually defined as the height at which the concentration of smoke has fallen to one-tenth of the concentration at the ground. Thus equating the exponential term to $\log_e 10$, then

$$2{\cdot}3 = \frac{3{\cdot}223z^{\frac{9}{7}}}{x}. \tag{4·8}$$

Thus at a distance of 100 m. downwind from the line source the height of a smoke cloud is

$$z = 9{\cdot}97 \text{ m., i.e. approximately } 10 \text{ m.},$$

and at a distance of 2000 m. $z = 100$ m.

Observations by the meteorological staff at a Ministry of Supply Experimental Station confirm these results. Thus they found that the concentration downwind over land with an adiabatic lapse rate varied as $x^{-0{\cdot}9}$, and the height of a smoke cloud at a distance of 100 m. from a line source was 10 m., irrespective of the wind speed.

O. G. Sutton (1932) has given a different expression for the distribution of matter in an air stream due to an infinite line source which is continuously emitting, i.e.

$$\chi = \frac{Bx^{-\frac{1}{1+p}}}{u}\exp\left\{-\frac{z^2}{C^2x^{\frac{2}{1+p}}}\right\},$$

where B and C are constant for any given value of p, and u does not vary with height.

Unlike (4·6) this expression was not obtained as a unique solution of the differential equation with appropriate boundary conditions, but was an expression selected arbitrarily to comply with certain boundary conditions.

It can be seen, however, that this expression is a solution of the two-dimensional equation of eddy diffusion

$$u\frac{\partial\chi}{\partial x} = \frac{\partial}{\partial z}\left\{K\frac{\partial\chi}{\partial z}\right\}$$

with the usual boundary conditions providing

$$K = \frac{C^2u}{2(1+p)}x^{-\frac{1-p}{1+p}}.$$

An expression for K, however, which is constant with height and which increases with distance downwind from the line source, does not accord with observation.

An exact expression for the distribution of water vapour in an air stream which after passing over a water surface passes over a land surface whose roughness parameter is the same as that of the sea may now readily be obtained.

For this problem x is measured downwind from the windward edge of the water surface whose length downwind is b (x is accordingly greater than b).

The evaporation at a distance λ downwind over the sea is, by (3·61),

$$A\lambda^{-\frac{m}{2m+1}}\chi_0, \tag{4·9}$$

where A is given by

$$A = \frac{(2m+1)^{\frac{1}{2m+1}}\,a^{\frac{m+1}{2m+1}}\,V_h h^{-m}}{\Gamma\left(\frac{m}{2m+1}\right)}. \tag{4·91}$$

The amount of vapour per unit volume at a height z and at a distance x due to a strip of water of length $d\lambda$ downwind evaporating at a rate given by (4·9) is, by (4·6),

$$\frac{A\lambda^{-\frac{m}{2m+1}}\chi_0\,d\lambda(x-\lambda)^{-\frac{m+1}{2m+1}}\exp\left\{-\frac{z^{2m+1}}{(2m+1)^2a(x-\lambda)}\right\}}{V_h h^{-m}(2m+1)^{\frac{1}{2m+1}}\,a^{\frac{m+1}{2m+1}}\,\Gamma\left(\frac{m+1}{2m+1}\right)} \tag{4·92}$$

The amount of vapour at a distance x downwind due to a sea track of length b is obtained by integrating this expression with respect to λ between the limits 0 and b is

$$\chi = \frac{\chi_0}{\Gamma\left(\frac{m}{2m+1}\right)\Gamma\left(\frac{m+1}{2m+1}\right)}\int_0^b \lambda^{-\frac{m}{2m+1}}(x-\lambda)^{-\frac{m+1}{2m+1}}\exp\left\{-\frac{z^{2m+1}}{(2m+1)^2a(x-\lambda)}\right\}d\lambda. \tag{4·93}$$ *

This form, which only applies to a special and rather artificial case, is not suitable for computation and is only given here for the sake of completeness. An approximate solution which is suitable for computation is given in the next section.

For $z = 0$ (4·9) becomes

$$\chi = \frac{\chi_0}{\Gamma\left(\frac{m}{2m+1}\right)\Gamma\left(\frac{m+1}{2m+1}\right)}\int_0^b \lambda^{-\frac{m}{2m+1}}(x-\lambda)^{-\frac{m+1}{2m+1}}d\lambda, \tag{4·94}$$

or writing $x = by$ and $\lambda = by\xi$, then

$$\begin{aligned}\chi &= \frac{\chi_0}{\Gamma\left(\frac{m}{2m+1}\right)\Gamma\left(\frac{m+1}{2m+1}\right)}\int_0^{\frac{1}{y}} \xi^{-\frac{m}{2m+1}}(1-\xi)^{-\frac{m+1}{2m+1}}d\xi,\\ &= \chi_0\left[1-\frac{1}{\Gamma\left(\frac{m}{2m+1}\right)\Gamma\left(\frac{m+1}{2m+1}\right)}\int_{\frac{1}{y}}^1 \xi^{-\frac{m}{2m+1}}(1-\xi)^{-\frac{m+1}{2m+1}}d\xi\right]\end{aligned} \tag{4·95}$$

For $m = \frac{1}{7}$

$$\chi = \chi_0\left[1-\frac{1}{\Gamma(\frac{1}{9})\,\Gamma(\frac{8}{9})}\int_{\frac{1}{y}}^1 \xi^{-\frac{1}{9}}(1-\xi)^{-\frac{8}{9}}d\xi\right], \tag{4·96}$$

---

* An alternative method of deriving this result has been given by W. G. L. Sutton (1943).

or writing $\mu = 1-\xi$

$$\chi = \chi_0\left[1-\frac{1}{\Gamma(\frac{1}{9})\,\Gamma(\frac{8}{9})}\int_0^{1-\frac{1}{y}}(1-\mu)^{-\frac{1}{9}}\mu^{-\frac{8}{9}}\,d\mu\right]. \tag{4·97}$$

If one now writes $y = 1+\epsilon$, where ϵ is small, then

$$\chi = \chi_0\left[1-\frac{\epsilon^{\frac{1}{9}}}{\frac{1}{9}\Gamma(\frac{1}{9})\,\Gamma(\frac{8}{9})}\right] \text{ approximately,} \tag{4·98}$$

$$\chi = \chi_0[1-0{\cdot}98\epsilon^{\frac{1}{9}}], \tag{4·99}$$

where $\epsilon = \dfrac{x-b}{b}$.

The ratio χ/χ_0 is given for values of ϵ from 0·001 to 0·10 in table 2.

TABLE 2

ϵ	0·001	0·002	0·003	0·004	0·005	0·006	0·007	0·008	0·009	0·010
χ/χ_0	0·545	0·507	0·486	0·469	0·456	0·445	0·435	0·427	0·419	0·412.
ϵ	0·01	0·02	0·03	0·04	0·05	0·06	0·07	0·08	0·09	1·0
χ/χ_0	0·412	0·365	0·337	0·314	0·297	0·283	0·270	0·260	0·250	0·240

As an illustration of the use of this table consider our earlier example in which air with an initial moisture content of 6·3 g./cu.m. flows over the Atlantic Ocean a distance of 5000 km., the surface temperature of the water being 49° F and the mean wind at a height of 10 m. being 7 m./sec.

After a further passage of 500 km. over land the moisture content at the ground should, from table 2, be given by

$$\mu-6{\cdot}3 = 0{\cdot}24[9{\cdot}09-6{\cdot}3]$$

or

$$\mu = 6{\cdot}96\text{ g.}$$

The mean moisture content of the air at Kew Observatory given by Bilham is 7·4 g./cu.m., which is not in good agreement with this calculated value.

Further consideration will be given to this problem in the following section, in which the degree of roughness of the land will be taken into account.

5. THE DISTRIBUTION OF WATER VAPOUR IN AIR WHICH AFTER PASSING OVER THE SEA FLOWS OVER THE LAND

From (3·7) the distribution of water vapour with height after a sea crossing of length b is to a high degree of approximation given by

$$\chi = \chi_0\left[1-\frac{z^m}{\Gamma\left(1+\frac{m}{2m+1}\right)[(2m+1)^2ab]^{\frac{m}{2m+1}}}\right], \tag{5·1}$$

where $a = mz_0^{2m}$, and similarly after a sea crossing of length b

$$-K\frac{\partial \chi}{\partial z} = D\chi_0, \tag{5·2}$$

where
$$D = \frac{maV_h h^{-m}}{\Gamma\left(1+\dfrac{m}{2m+1}\right)[(2m+1)^2ab]^{\frac{m}{2m+1}}} = \text{a constant.} \tag{5·3}$$

Now in considering the subsequent distribution of water vapour over land the equation stating the advection and diffusion balance is as before

$$z^m\frac{\partial \chi}{\partial x} = a_1\frac{\partial}{\partial z}\left\{z^{1-m}\frac{\partial \chi}{\partial z}\right\},$$

where $a_1 = mz_{01}^{2m}$, where z_{01} is a measure of the roughness of the land. (In this equation x is measured downwind from the windward edge of the land surface.)

If
$$S = -a_1V_h h^{-m}z^{1-m}\frac{\partial \chi}{\partial z},$$

then
$$\frac{\partial S}{\partial x} = a_1 z^{1-m}\frac{\partial}{\partial z}\left\{z^{-m}\frac{\partial S}{\partial z}\right\}, \tag{5·4}$$

and a solution of (5·4) is required subject to the boundary conditions

(1) $S \to D\chi_0$ when $x \to 0$,

(2) $S \to 0$ when $z \to 0,\ x > 0$.

(It is assumed that the land is a dry surface impermeable to water vapour.)

As in § 3, an examination of the dimensions of the terms of (5·4) suggests that it should be possible to find a solution $S = f(\theta)$, where

$$\theta = \frac{z^{2m+1}}{(2m+1)^2 a_1 x}. \tag{5·41}$$

Denoting differentiation with respect to θ by dashes, such a solution must satisfy

$$\theta S'' + \left(\theta + \frac{m}{2m+1}\right)S' = 0, \tag{5·5}$$

and hence
$$S = \frac{D\chi_0\displaystyle\int_0^\theta e^{-\theta}\theta^{-\frac{m}{2m+1}}\,d\theta}{\Gamma\left(\dfrac{m+1}{2m+1}\right)} \tag{5·6}$$

clearly satisfies all the conditions of (5·4),

but
$$S = -a_1V_h h^{-m}z^{1-m}\frac{\partial \chi}{\partial z},$$

therefore
$$\frac{\partial \chi}{\partial z} = -\frac{D\chi_0 z^{m-1}}{a_1 V_h h^{-m}\Gamma\left(\frac{m+1}{2m+1}\right)}\int_0^\theta e^{-\theta}\theta^{-\frac{m}{2m+1}}\,d\theta,$$

and hence, integrating by parts, it is seen that

$$\chi = \chi_0\left\{1 - \frac{D}{ma_1 V_h h^{-m}\Gamma\left(\frac{m+1}{2m+1}\right)}\left[z^m\int_0^\theta e^{-\theta}\theta^{-\frac{m}{2m+1}}\,d\theta + e^{-\theta}[(2m+1)^2 a_1 x]^{\frac{m}{2m+1}}\right]\right\}$$

$$= \chi_0\left\{1 - \frac{az^m\int_0^\theta e^{-\theta}\theta^{-\frac{m}{2m+1}}\,d\theta + e^{-\theta}a[(2m+1)^2 a_1 x]^{\frac{m}{2m+1}}}{a_1\Gamma\left(\frac{m+1}{2m+1}\right)\Gamma\left(1+\frac{m}{2m+1}\right)[(2m+1)^2 ab]^{\frac{m}{2m+1}}}\right\}$$

$$= \chi_0\left\{1 - \left(\frac{a}{a_1}\right)^{\frac{m+1}{2m+1}}\left(\frac{x}{b}\right)^{\frac{m}{2m+1}}\left[\frac{e^{-\theta} + \theta^{\frac{m}{2m+1}}\int_0^\theta e^{-\theta}\theta^{-\frac{m}{2m+1}}\,d\theta}{\Gamma\left(1+\frac{m}{2m+1}\right)\Gamma\left(\frac{m+1}{2m+1}\right)}\right]\right\} \tag{5·7}$$

is the required solution.

Or writing $m = \frac{1}{7}$, then

$$\chi = \chi_0\left\{1 - 0{\cdot}98\left(\frac{x}{b}\right)^{\frac{1}{9}}\left(\frac{z_0}{z_{01}}\right)^{\frac{16}{63}}\left[e^{-\theta} + \theta^{\frac{1}{9}}\int_0^\theta e^{-\theta}\theta^{-\frac{1}{9}}\,d\theta\right]\right\}. \tag{5·71}$$

For small values of θ (5·71) becomes

$$\chi = \chi_0\left\{1 - 0{\cdot}98\left(\frac{x}{b}\right)^{\frac{1}{9}}\left(\frac{z_0}{z_{01}}\right)^{\frac{16}{63}}\left(1+\frac{\theta}{8}\right)\right\}, \tag{5·8}$$

which is suitable for computation.

The problem of the previous section may now be reconsidered, in which air with an initial moisture content of 6·3 g./cu.m. flows over the Atlantic Ocean a distance of 5000 km., the surface temperature of the water being 49° F, and the mean wind speed at a height of 10 m. being 7 m./sec.

On reaching the English coast the amount of water vapour per cubic metre of air at a height of 1·2 m. above the surface has been shown to be 8·3 g./cu.m.

After a further passage of 500 km. over the land the amount of water vapour per cubic metre of air at a height of 1·2 m. above the surface would be, from (5·8),

$$\mu - 6{\cdot}3 = (9{\cdot}09 - 6{\cdot}3)\left[1 - \left(\frac{z_0}{z_{01}}\right)^{\frac{16}{63}}\left(\frac{1}{10}\right)^{\frac{1}{9}} \times 0{\cdot}98\left(1 + \frac{(120)^{\frac{9}{7}}.343}{5\times 10^7\times 8\times 81\times z_{01}^{\frac{2}{7}}}\right)\right].$$

With $z_{01} = 2{\cdot}6$ cm., which is the value derived from Taylor's observations of the shearing stress,

$$\mu = 7{\cdot}4 \text{ g./cu.m.,} \tag{5·81}$$

which is in excellent agreement with the value of the mean moisture content at Kew Observatory given by Bilham.

I should like to express my thanks to Mr E. Gold for his encouraging interest in this work and to acknowledge my debt to the late Dr F. J. W. Whipple for his help with § 3 of this paper.

REFERENCES

Bilham 1938 *Climate of the British Isles.*
Brunt 1939 *Physical and dynamical meteorology*, 2nd ed.
Carruthers 1943 *Quart. J.R. Met. Soc.* **69**, 289.
Ertel 1933 *Met. Z.* **50**, 386.
Goldstein 1938 *Modern developments in fluid dynamics.*
Sutton, O. G. 1932 *Proc. Roy. Soc.* A, **135**, 668.
Sutton, O. G. 1934 *Proc. Roy. Soc.* A, **146**, 701.
Sutton, W. G. L. 1943 *Proc. Roy. Soc.* A, **182**, 48.
Sverdrup 1936 *Geofys. Publ.* **11**, no. 7.
Wust 1920 *Geogr. Ann., Stockh.*, 1920, no. 3.

Contributions to the theory of chromatography

BY E. GLÜCKAUF

(*Communicated by E. J. Bowen, M.A., F.R.S.—Received* 6 *June* 1945)

A complete analysis has been given of the process of chromatographic separation for two solutes, the adsorption of which follows a Langmuir isotherm. Development of the bands, both with pure solvent and with a solvent containing another solute, has been discussed. Very simple conditions were found to hold for both the volume of solvent and the amount of adsorbent required for complete separation of the two solutes. Provided that the concentrations used are not too low—which is unfavourable—the influence of the initial concentrations is very small. For substances difficult to separate, the minimum amount of adsorbent X_0 (in g.) is proportional to the quantities to be separated (m_1+m_2) (in mol.), and inversely to the saturation capacity of the adsorbent $(1/\beta)$ (in mol./g. adsorbent) and to the square of the relative difference of the adsorption coefficients (a_1 and a_2) of the solutes

$$X_0 = \text{approximately } \beta(m_1+m_2)/(a_2/a_1-1)^2.$$

The accurate equations are given in the paper.

The main contributions to the theory of chromatography have been made by Wilson (1940), by de Vault (1943) and by Weiss (1943).

Wilson was the first to develop the differential equations governing the process of chromatography, but, as both de Vault and Weiss have shown, these equations are capable of more than one solution. The solution found by Wilson gives correct results only, when the adsorbed amount is strictly proportional to the concentration of solute in the solvent.

Weiss has shown how the properties and positions of a band of a single solute change with progressing chromatographic development, but the case of two or more solutes has so far not been treated beyond the stage of the differential equations. Wilson's deductions have left the impression that the length of column and the amount of pure solvent required for the separation of two substances are inversely proportional to $(a_2-a_1)/a_1$, where a_2 and a_1 are the adsorption coefficients of the two solutes ($a_2>a_1$). As will be shown in this paper this result is by far too optimistic, especially in cases where the differences of adsorption intensities are only slight.

In the following discussion the adsorption isotherms have been assumed to be of the Langmuir type, which is the only isotherm showing the mutual influence of the concentrations on the adsorption isotherms of the components of the system. During formation and development of the bands, the equations of motion for the different parts of the bands have often a form similar to those given by Weiss for a single solute, and frequent comparison will be made with his equations (marked W.). In the following, the subscripts 1 and 2 mark the two bands to be separated while the superscript 0 refers to the original concentrations; letter subscripts refer to the positions in the band as shown in the diagrams. It is assumed that the subscript 1 refers to the faster moving solute.

1. THE CASE OF TWO SOLUTES SEPARATED BY NORMAL DEVELOPMENT

As in the case of a single solute, the bands formed originally are of constant intensity throughout, except where they overlap, and the rate of movement of the forward edge of the constant part is given by

$$x_u = (v+v^0) . c_2^0/(f_{2(c^0{}_1c^0{}_2)}+\alpha c_2^0) \tag{1}$$

(see figure 1*b*, also compare with equation W. (11) and de Vault, equation 9), where

x is the distance of the particular point from the top of the column (best expressed in grams of adsorbent),

v^0 is the original volume of solvent in which the solutes of concentration c_1^0 and c_2^0 were contained (v^0 in c.c.),

v is the amount of pure solvent used for developing the chromatogram (in c.c.),

c concentration of solute in mol./c.c.,

α is the pore volume of the column per g. of adsorbent.

$f_{2(c_1c_2)} = q_2$ is the equilibrium amount of substance 2 as function of the concentrations c_1 and c_2, adsorbed on a gram of adsorbing material (q in mol./g.).

$(f_{2(c_1c_2)}+\alpha c_2)$ thus represents the amount of solute II contained in the space filled by 1 g. of adsorbent, if the local concentrations of solutes in the solvent are c_1 and c_2. The corresponding amount of solute I is $(f_{1(c_1c_2)}+\alpha c_1)$. When the development proceeds, regions of varying concentration appear in the chromatogram, as in the case of a single solute. In these regions, the movement of any points of given concentrations proceeds according to

$$x = v/(\alpha+f'_{1(c_1c_2)}) = v/(\alpha+f'_{2(c_1c_2)}), \tag{2}$$

where $f'_{1(c_1c_2)} = dq_1/dc_1$ and vice versa.* De Vault (1943, equations 19, 20 and 23) has shown that equation (2) results directly from the consideration that the difference between the amounts of solutes entering and leaving a small section of the column must necessarily find their balance in the change of the amount of solutes contained in that section. As pointed out by de Vault, it is not possible to write down a simple solution for the case of multiple solutes, as the dependence of c_1 and c_2 can normally not be expressed in general terms. However, equations (1) and (2), together with the Langmuir adsorption isotherms (3, 4) of the two substances, permit of a general solution of all questions affecting the development of the chromatogram. Other isotherms, e.g. that of Freundlich, are quite unsuitable, because the mutual influence of two solvents cannot be formally expressed.

The Langmuir adsorption isotherms in the case of two solutes are

$$f_{1(c_1c_2)} = q_1 = a_1c_1/(1+a_1\beta_1c_1+a_2\beta_2c_2), \tag{3}$$

$$f_{2(c_1c_2)} = q_2 = a_2c_2/(1+a_1\beta_1c_1+a_2\beta_2c_2), \tag{4}$$

where a = adsorption coefficients (dimension c.c./g. adsorbent) $a_2 > a_1$,

β = mol. capacity, i.e. the amount of adsorbent required by 1 mol. of substance if fully covering the adsorbing surface in a mono-molecular layer (dimension g./mol.). For simplicity use is also made of the letters,

$b = a\beta$,

$\delta = (a_2-a_1)/a_1$.

It follows from equations (3) and (4) that

$$f'_{1(c_1c_2)} = a_1(1+b_2c_2-b_2c_1\,.\,dc_2/dc_1)/(1+b_1c_1+b_2c_2)^2, \tag{3a}$$

$$f'_{2(c_2c_2)} = a_2(1+b_1c_1-b_1c_2\,.\,dc_1/dc_2)/(1+b_1c_1+b_2c_2)^2. \tag{4a}$$

From the equality of f'_1 and f'_2 (see equation (2)) at any given point or stage of the chromatogram it is possible to calculate the relation which must exist between any two coexistent concentrations of the two solutes.

Equating (3a) and (4a) results in the differential equation

$$a_2-a_1+a_2b_1c_1-a_1b_2c_2+a_1b_2c_1\,.\,dc_2/dc_1-a_2b_1c_2\,.\,dc_1/dc_2 = 0, \tag{5}$$

the only general solution of which is

$$a_2b_1c_1 = a_1b_2c_2\lambda-(a_2-a_1)\,\lambda/(1+\lambda), \tag{6}$$

where λ is an integration constant which can be calculated from the boundary conditions applying, e.g.

$$c_1 = c_1^0, \quad \text{when} \quad c_2 = c_2^0, \tag{7}$$

---

* $f'_{2(-,-)}$ always stands for the complete differential $df_{2(c_1c_2)}/dc_2$, the contents of the bracket indicating the values of c_1 and c_2, if these are defined; e.g. $f'_{2(0,c_{20})} = [df_{2(c_1c_2)}/dc_2]_{c_1=0,\,c_2=c_{20}}$. If one of the concentrations is omitted, e.g. $f'_{2(c_{20})}$, this means that the missing concentration is zero throughout, thus: $f'_{2(c_{20})} = [\delta f_{2(c_1c_2)}/\delta c_2]_{c_1=0,\,c_2=c_{20}}$. The same applies to the f'_1-functions.

which applies to the development of the chromatogram from its beginning if the same solvent is used. Equation (6) is linear for c_1 and c_2 but quadratic for λ:

$$\lambda^2+\lambda(1-a_2b_1c_1^0/a_1b_2c_2^0-(a_2-a_1)/a_1b_2c_2^0)-a_2b_1c_1^0/a_1b_2c_2^0=0. \tag{8}$$

The positive root λ_1 (or shortly λ) of equation (8) is shown in figure 2 as function of $a_2b_1c_1^0/a_1b_2c_2^0$ and of $\delta/b_2c_2^0$, and λ_1 applies to all cases where dc_1/dc_2 is positive. The negative root (λ_2) can be used only for a single point at the frontal edge of the

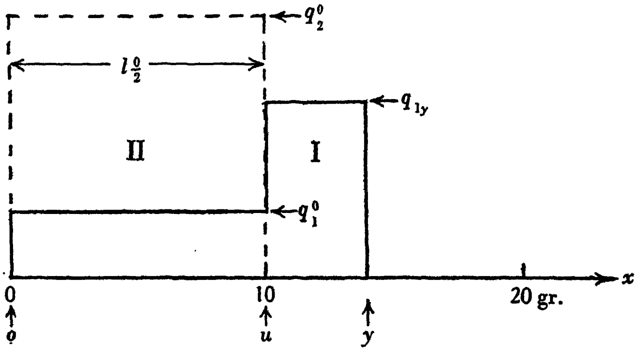


FIGURE 1a. $v_0 = 100$ c.c., $v = 0$.

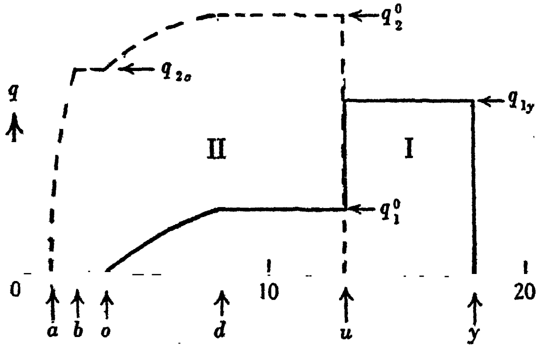


FIGURE 1b. $v = 30$ c.c.

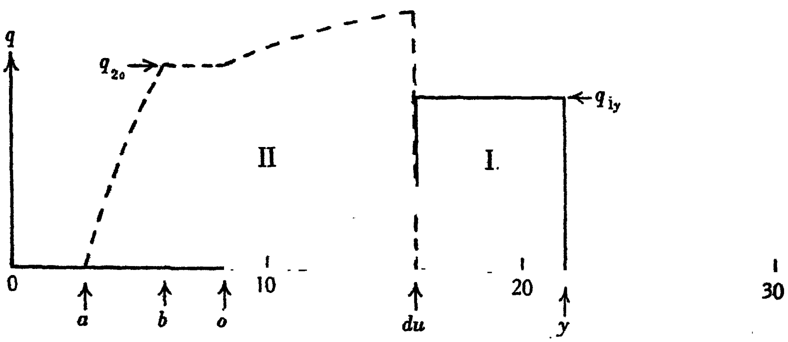


FIGURE 1c. $v_{du} = 57$ c.c.

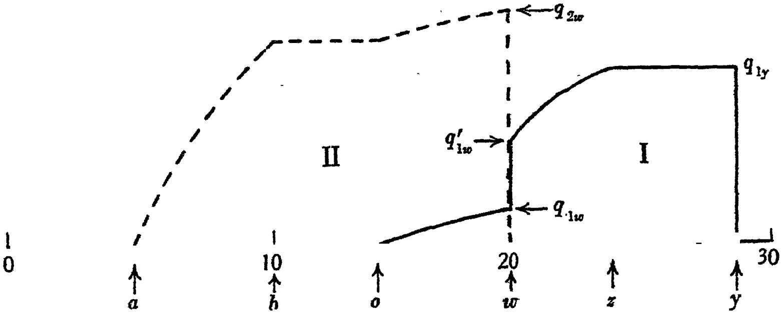


FIGURE 1*d*. $v = 100$ c c.

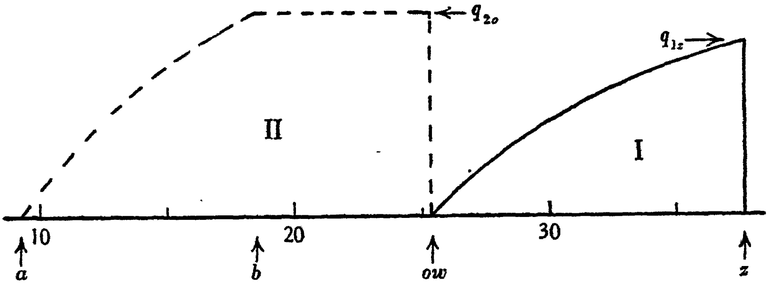


FIGURE 1*e*. $V = 174$ c.c.

FIGURE 1*a*–*e*. Development with pure solvent of a band of two solutes ($\delta = 1$). (*a*) Band before development. (*b*) Initial development; original concentrations still preserved in the forward part of the band. (*c*) Mixed band fully developed, but separated part of band I still has its original concentration. (*d*) Further development and spreading of the bands. (*e*) Separation completed; development of band I is also completed. Apparent equality of q_{1z} and q_{1v} is only a coincidence.

slower band II because, as shown by de Vault, the general solution for the front boundary of the differential equations on which equation (2) is based is unsuitable for this type of adsorption isotherm and has to be replaced by a discontinuity. λ_2 does, however, represent correctly the sudden change of c_1 at this discontinuity. No table is given for λ_2, as its values follow directly from λ_1 and equation (8) as

$$\lambda_1 \lambda_2 = -a_2 b_1 c_1^0 / a_1 b_2 c_2^0. \tag{9}$$

It can be seen from figure 2 that for substances very difficult to separate ($\delta \ll 1$), λ_1 is approximately c_1^0/c_2^0, (usually $\beta_1 = \beta_2$), while in the event that δ is large or c_2^0 is very small, $\lambda_1 = \delta/b_2 c_2^0$. The conditions in the variable parts of the chromatogram can now be given as functions of v, x and c_1 or c_2 by substituting the differentials and one of the variable concentrations of (3*a*) and (4*a*) from equation (6) and using the f'-functions in equation (2). If α is neglected, this results in

$$x = v(1 + B_1 c_1)^2 / A_1 = v / F_1'(c_1), \tag{10}$$

$$x = v(1 + B_2 c_2)^2 / A_2 = v / F_2'(c_2) \tag{11}$$

(compare equation W. (26) for single solute), where

$$A_1 = a_1(1+\lambda)/(1+\lambda+\delta),$$

$$B_1 = b_1(1+\lambda)/\lambda,$$

$$A_2 = a_2(1+\delta)(1+\lambda)/(1+\lambda+\delta),$$

$$B_2 = b_2(1+\lambda).$$

It can be shown easily that the equations (10) and (11) fulfil the fundamental equations for mass conservation (see de Vault equation (19))

$$\delta c_i/\delta x + \delta q_i/\delta v = 0.$$

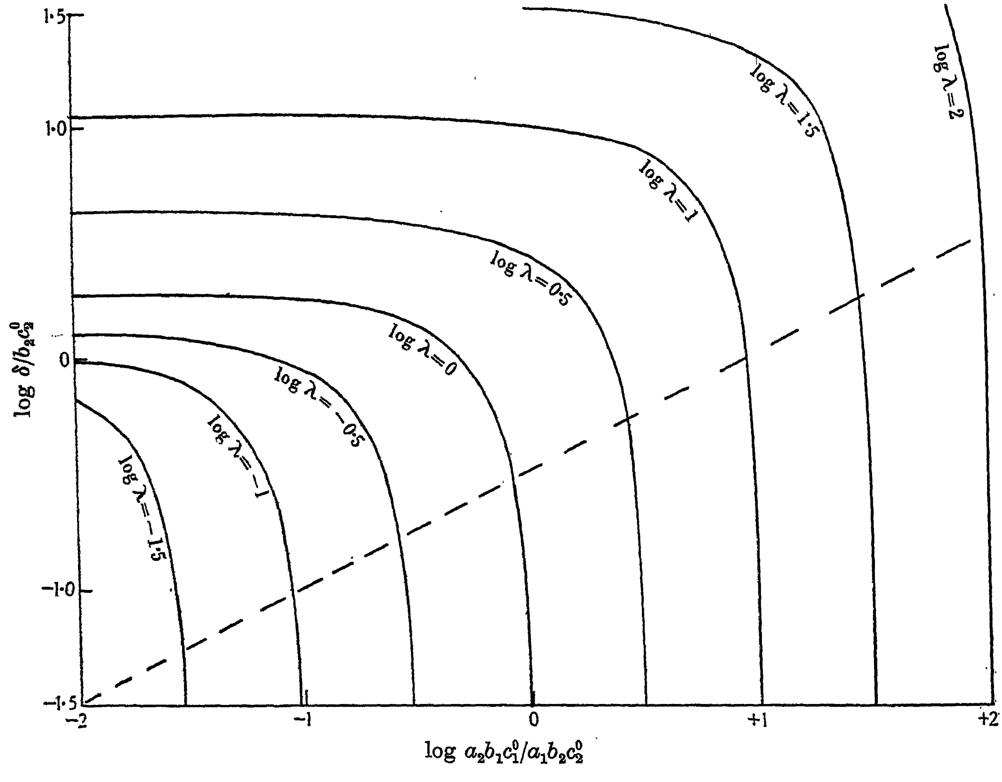


Figure 2. Values of λ as function of $\delta/b_2c_2^0$ and $a_2b_1c_1^0/a_1b_2c_2^0$. Optimum separation conditions which would require a minimum of both solvent and adsorbent are obtained below the dotted line.

2. Formation of band

If an amount v^0 of a solvent containing two solutes I and II of the concentrations c_1^0 and c_2^0 (where the adsorption intensities are $a_2 > a_1$) is poured on to the chromatographic column, the initial band has the form as shown, e.g. in figure 1*a*. The properties of this band with the exception of c_{1y} can be derived immediately from equations (1), (3) and (4):

$$x_u = l_2^0 = v^0 c_2^0 / f_{2(c^0_1 c^0_2)} = v_0(1+b_1 c_1^0 + b_2 c_2^0)/a_2, \tag{12}$$

$$\left.\begin{aligned} q_1^0 &= f_{1(c^0_1 c^0_2)} = a_1 c_1^0/(1+b_1 c_1^0 + b_2 c_2^0), \\ q_2^0 &= f_{2(c^0_1 c^0_2)} = a_2 c_2^0/(1+b_1 c_1^0 + b_2 c_2^0). \end{aligned}\right\} \tag{13a, b}$$

If m_1 and m_2 are the total quantities of the solutes (in mol.), the amount of solute I within the length l_2^0 is

$$m_{1,o-u} = m_2 q_1^0/q_2^0 = m_1 a_1/a_2, \tag{14}$$

while the amount of solute I separated off already during band formation is

$$m_{1,u-y} = m_1(a_2 - a_1)/a_2. \tag{14a}$$

As predicted already by de Vault the concentration of the less adsorbed solute (I) rises sharply at the discontinuity at x_u to a value which follows from equation (6) for $c_2 = 0$

$$c_{1y} = -(a_2 - a_1)\lambda_2/a_2 b_1(1+\lambda_2). \tag{15}$$

Replacing λ_2 from equation (9) and then c_2 by means of equation (6), this leads to

$$c_{1y} = c_1^0(1+\lambda_1)/\lambda_1. \tag{15a}$$

As can be seen from figure 2, c_{1y} must have values between c_1^0 and $(c_1^0 + c_2^0)$. It also follows from equation (3) that

$$q_{1y} = a_1 c_{1y}/(1+b_1 c_{1y}), \tag{16}$$

and from equation (1) applied to this case, that the position of the forward edge is

$$x_y = v^0(1+b_1 c_{1y})/a_1. \tag{17}$$

3. DEVELOPMENT OF THE CHROMATOGRAM

If, after formation of the original chromatogram, a further amount v of solvent is poured on the column, the bands, dissolving from the top end and travelling down the column, undergo certain changes, some of which are similar to those occurring with a single solute (see figure 1*b*–*d*). The first stage in the development is shown in figure 1*b*. The concentrations in the most forward parts of the column remain unchanged; then follows a part in which both concentrations decrease, the less adsorbed one decreasing to zero; then comes a constant region of pure substance II, followed by a variable region in which this concentration, too, diminishes to zero. Figure 1*c* shows the end of the first stage, with band II completely developed.

The concentration of the constant region of pure substance II follows directly from equation (6) for $c_1 = 0$:

$$c_{2o} = \delta/b_2(1+\lambda). \tag{18}$$

The other conditions are

$$x_a = v/f'_{2(c_2=0)} = v/a_2 \quad \text{(see equation W. (23))}, \tag{19}$$

$$x_b = v/f'_{2(c_{2o})} = v(1+b_2c_{2o})^2/a_2 \quad \text{(see equation W. (24))}, \tag{20}$$

$$x_o = v/F'_{1(0)}, \quad \text{(see equation 10)} \tag{21}$$

$$x_d = v/F'_{1(c^0{}_1)} = v/F'_{2(c^0{}_2)}, \quad \text{(see equations 10, 11)} \tag{22}$$

$$x_u = (v+v^0)\,c_2^0/f_{2(c^0{}_1c^0{}_2)} = (v+v^0)\,(1+b_1c_1^0+b_2c_2^0)/a_2, \tag{23}$$

$$x_y = (v+v^0)\,c_{1y}/f_{1(c_{1y})} = (v+v^0)\,(1+b_1c_{1y})/a_1. \tag{24}$$

It is worth while noting that x_b can never catch up with x_o, as $x_o/x_b = 1+\dfrac{\delta\lambda}{1+\lambda+\delta}$. It is also interesting that the pure tail fraction between x_a and x_o contains the same amount of material, as if the curve of the variable part between x_d and x_o would take its uninterrupted course.

The content of the pure tail fraction $m_{2(a-o)}$ at any stage of the separation process follows from a very simple consideration:

If the separated band containing $m_{2(a-o)}$ were not constantly fed from the tail end of the mixed band, its forward edge would travel at the rate

$$\frac{\Delta x'_{2o}}{\Delta v} = \frac{c_{2o}}{\alpha c_{2o}+f_{2(0,c_{2o})}} \quad \text{(see equation (1))}.$$

Actually x_{2o} travels according to

$$\frac{\Delta x_{2o}}{\Delta v} = \frac{1}{\alpha+f'_{1(0,c_{2o})}} \quad \text{(see equation (2) or (21))}.$$

In order to maintain a continuous chromatogram, the amount of solute II introduced by feeding from the mixed band must be

$$m_2 = (\Delta x_{2o}-\Delta x'_{2o})\,(\alpha c_{2o}+f_{2(0,c_{2o})}),$$

from which follows

$$m_{2(a-o)} = \frac{f_{2(0,c_{2o})}-c_{2o}f'_{1(0,c_{2o})}}{\alpha+f'_{1(0,c_{2o})}}\,v. \tag{25}$$

Neglecting α and introducing c_{2o} from equation (18) leads to the very simple condition

$$m_{2(a-o)} = v\delta c_{2o} = v\delta^2/b_2(1+\lambda). \tag{25a}$$

The effect of this equation on the question of complete separation will be shown later.

The contents of the pure head fraction at this stage, i.e. as long as its concentration is constant, is

$$m_{1,u-y} = (v+v^0)\,c_1^0(a_2-a_1)/a_2. \tag{26}$$

These conditions persist until x_d overtakes x_u, when development of band II is complete (see figure 1*c*). This stage is reached when an amount v_{du} of solvent has been used, which can be obtained by equating equations (22*a*) and (23):

$$v_{du} = v^0/b_2c_2^0(1+\lambda) = m_2/b_2(c_2^0)^2\,(1+\lambda). \tag{27}$$

After band II has been fully developed, its frontal concentration c_{2w} (see figure 1*d*) starts to decrease according to

$$c_{2w} = \sqrt{(m_2/B_2 v)} = \sqrt{(m_2/b_2 v(1+\lambda))}, \tag{28}$$

in complete analogy to the conditions of a single solute (see equation W. 32)). The movement of the frontal edge of band II is given by

$$\sqrt{x_w} = \sqrt{(v/A_2)} + \sqrt{(m_2 B_2/A_2)} \tag{29}$$

(compare equation W. (31)).

The concentration of band I at the forward edge of band II (c_{1w}) follows from equations (28) and (6):

$$c_{1w} = \lambda(a_1 b_2/a_2 b_1)\sqrt{(m_2/b_2 v(1+\lambda))} - \lambda(a_2-a_1)/a_2 b_1(1+\lambda). \tag{30}$$

The concentration c'_{1w} at the other side of the discontinuity also follows from equation (6) for $c'_{2w} = 0$:

$$c'_{1w} = -\lambda'_2(a_2-a_1)/(1+\lambda'_2)\,a_2 b_1. \tag{31}$$

The boundary condition for determining the integration constant λ'_2 is in this case no longer $c_1 = c_1^0$ when $c_2 = c_2^0$, but $c_1 = c_{1w}$ when $c_2 = c_{2w}$. This gives the same *value of* λ_1, *but we obtain now*

$$\lambda'_2 = -a_2 b_1 c_{1w}/a_1 b_2 c_{2w}. \tag{32}$$

Substituting λ'_2 into equation (31) results in

$$c'_{1w} = c_{1w}(1+\lambda_1)/\lambda_1. \tag{33}$$

After band II has been fully developed, the head fraction of pure solute I, too, undergoes development. Being a pure solute, a point of given concentration c_1 moves forward according to

$$\Delta x = \Delta v(1+b_1 c_1)^2/a_1 \tag{34}$$

(see equation (2)), and the boundary conditions are that this movement starts at a point where $c_1 = c'_{1w}$, which occurs according to equation (33) and (30) at

$$v_w = a_1^2 b_2 m_2(1+\lambda)/(a_2 b_1 c'_{1w} + a_2 - a_1)^2 \tag{35}$$

and

$$x_w = v(1+b_1 c'_{1y})^2/A_1, \tag{36}$$

which latter can be shown to be the locus of all points c'_{1w}. Combining equations (30), (34), (35) and (36) one obtains for the movement of points between x_w and x_z (where x_z is the most forward part of the head fraction, which has been reached by the development)

$$x = (v-v_w)(1+b_1 c_1)^2/a_1 + x_w,$$

$$x = v(1+b_1 c_1)^2/a_1 + \delta m_2 b_2(1+b_1 c_1)^2/a_1(\delta+b_2 c_1)^2. \tag{37}$$

While equation (37) permits of the calculation of c_1 at any given x and v, it seems no longer possible to obtain a simple equation for the end-point x_z and the frontal concentration c_{1z}, though these facts are well defined by the condition of mass conservation

$$\int_{x_0}^{x_w} q_1\,dx + \int_{x_w}^{x_t} q_1\,dx = m_1,$$

and thus can be constructed in any given numerical case (as has been done in figure 1 *e*).

In some instances, e.g. when only partial separation is required, it may be desirable to know the amount of solute in a particular region for the mixed band.

The amount of solute I between x_p and x_q is

$$m_{1(p-q)} = \int_{x_p}^{x_q} a_1 c/(1+b_1c_1+b_2c_2)\,dc_1.$$

Replacing c_2 from equation (6) and dc_1 by dx from equation (10) this leads to

$$m_{1(p-q)} = 2vB_1\int_{x_p}^{x_q} c_1\,dc_1 = B_1v(c_{1q}^2 - c_{1p}^2). \tag{38}$$

Similarly,

$$m_{2(p-q)} = B_2v(c_{2q}^2 - c_{2p}^2). \tag{39}$$

4. COMPLETE SEPARATION

Development continues to proceed until the bands separate whereby the frontal concentration in band I becomes continually smaller. The stage of complete separation can be obtained from a number of previous equations, e.g.

(1) from equation (25) by making $m_{2,a-o} = m_2$,
(2) by equating x_o (see equation (21)) and x_w (see equation (29)),
(3) by making c_{1w} (equation (30)) or c'_{1w} (equation (33)) zero.

The result is in all cases that the volume V required for complete separation is

$$V = m_2b_2(1+\lambda)/\delta^2. \tag{40}$$

The point of separation of the two bands has a distance from the top of the column which is given by

$$X_o = V/A_1 = m_2b_2(1+\lambda+\delta)/a_1\delta^2. \tag{41}$$

The tail end of band II is at

$$X_a = V/a_2 = m_2b_2(1+\lambda)/a_2\delta^2. \tag{42}$$

The length of band II at the time of separation is thus

$$L_2 = m_2b_2(2+\lambda+\delta)/a_2\delta. \tag{43}$$

In the case of extreme dilution, when $\delta/b_2c_2^0 \gg 1$, λ approximately equals $\delta/b_2c_2^0$, and we obtain

$$V = m_2/\delta c_2^0 = v^0/\delta, \tag{40a}$$

and, assuming that $\delta < 1$,

$$X_a = m_2/\delta a_2 c_2^0 = m_2/\delta q_2^0 = 1_2^0/\delta. \tag{42a}$$

The length of band II at the point of separation is then

$$L_2 = m_2/a_2 c_2^0 = 1_2^0, \tag{43a}$$

i.e. it has remained unchanged. These are the equations obtained by Wilson (1941) for non-developing bands. At the same time, it can be seen from equations (40)–(42) that this is the most unfavourable condition for a chromatographic separation, as it requires an unnecessarily large amount of both solvent and adsorbent.

The most favourable conditions obviously occur when λ is small. The smallest values of λ usually occur when $\delta/b_2 c_2^0 < 1$, i.e. at moderately high adsorption intensities.* As can be seen from table 2, optimum conditions can be defined as

$$b_2 c_2^0 > \delta\sqrt{(10c_2^0/c_1^0)} \quad \text{or} \quad b_2\sqrt{(c_1^0 c_2^0)} > 3\delta,$$

which condition obtains below the dotted line in figure 2.

Under these optimum conditions, where λ = approx. m_1/m_2, equations (40) and (41) change into

$$V = b_2(m_1+m_2)/\delta^2, \tag{40b}$$

and

$$X_o = \beta(m_1+m_2)(1+\delta)/\delta^2, \tag{41b}$$

which means that within this region both the amount of solvent and the weight of adsorbent required for the complete separation are independent of the original concentrations used, and depend only on the total quantity of solutes to be separated.

In this connexion it may be noted that X_o is the necessary amount of adsorbent in the chromatographic tube up to the point of separation and that consequently, as long as the effect of diffusion is neglected, a short column of large diameter will serve equally well as a long and narrow one.

5. PORE SPACE OF THE ADSORBENT

In most of the preceding calculations the question of pore space has not been considered for reasons of simplicity. It may, however, be worth while to consider its effect on the volume of solvent and the length of column required for complete separation. It follows from equation (25) that

$$V_\alpha = m_2(\alpha + f'_{1(0,c_{2o})})/(f_{2(0,c_{2o})} - c_{2o} f'_{1(0,c_{2o})}); \tag{44}$$

eliminating c_{2o}, this results in

$$V_\alpha = V\left(\frac{\alpha(1+\lambda+\delta)}{a_1(1+\lambda)} + 1\right). \tag{45}$$

Further, it follows from equations (44) and (2) that

$$X_\alpha = m_2(f_{2(0,c_{2o})} - c_{2o} f'_{1(0,c_{2o})}) = X_o. \tag{46}$$

This shows that while a correction is required for the volume of solvent, the amount of adsorbent is not influenced by the pore space of the chromatographic column.

---

* According to the theoretical interpretation of the Langmuir equation, $b_2 c_2$ represents the ratio of active surface covered by solute II to surface not covered by this solute.

6. Development with a solvent containing another solute

A. *Development of band of single solute*

The subscript 3 is used here for the properties of the developing solute, and the index 2 for the band. According to the concentrations employed and to the adsorption intensities involved, different conditions arise as regards the effect of the developing solute.

Case 1. $f'_{2(0)} \leqslant f_{3(c^0_3)}/c^0_3$ or $a_2 \leqslant a_3/(1+b_3c^0_3)$.

In this case the developing solution acts like a pure solvent, as all the solute III is adsorbed before it makes contact with band II. This is obvious, as the forward edge of solute III advances according to

$$x = vc^0_3/f_{3(c^0_3)} \tag{47}$$

while the rear of band II progresses with

$$x = v/f'_{2(0)}. \tag{48}$$

Case 2. $f_{3(c^0_3)}/c^0_3 < f'_{2(0)} \leqslant f'_{3(0)}$ or $a_3/(1+b_3c^0_3) < a_2 \leqslant a_3$.

In this case, band II undergoes only partial development (see figure 3*a*). The concentration c_{2o} is given by the condition

$$f_{3(c^0_3)}/c^0_3 = f_{2(c_{2o})}/c_{2o}, \tag{49}$$

in the case of the Langmuir isotherm by

$$c_{2o} = c^0_3a_2b_3/a_3b_2 + (a_2-a_3)/a_3b_2.$$

With progressing development band II eventually assumes the form shown by the dotted line. This condition also applies, if $c^0_2 < c_{20}$. In this case a discontinuous development results (see figure 3*b*) which, after completion, leaves band II in the form shown by the dotted line. Once the development of band II has been completed, no further change in the length or concentration of band II occurs.

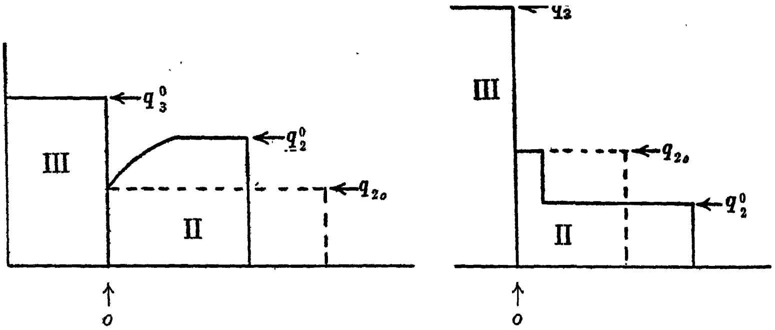


FIGURE 3*a*, *b*. Development of a band containing one solute with a solvent containing a solute which is more strongly adsorbed. (*a*) Small concentration of developing solute. (*b*) Large concentration of developing solute.

Case 3. $f'_{2(0)} > f'_{3(0)}$ or $a_2 > a_3$.

In this case development results in the formation of a mixed band of variable concentration, the conditions of which can only be calculated for the Langmuir equation.

In the variable part equations (2)–(6) apply again, but the boundary condition for the determination of λ is in this case

$$c_2 = 0 \quad \text{when} \quad c_3 = c_3^0.$$

By applying this condition to equation (6),

$$\lambda = \frac{-a_2 b_3 c_3^0}{a_2 - a_3 + a_2 b_3 c_3^0}.$$

Thus

$$c_3 = c_3^0\left(1 - \frac{a_3 b_2 c_2}{a_2 - a_3 + a_2 b_3 c_3^0}\right). \tag{50}$$

The conditions in the variable part can now be given as functions of v, x and c_2, by substituting c_3 in the equation of the type (4*a*) from equation (50), and by using the f'-function in equation (2). This results in

$$x = v(1 + B_2 c_2)^2 / A_2, \tag{51}$$

where

$$A_2 = a_2/(1 + b_3 c_3^0) \tag{51a}$$

and

$$B_2 = b_2(a_2 - a_3)/(a_2 - a_3 + a_2 b_3 c_3^0). \tag{51b}$$

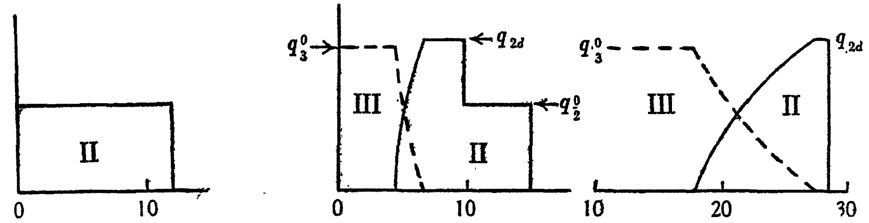


FIGURE 4*a*. $v_0 = 200, v = 0$. FIGURE 4*b*. $v = 50$. FIGURE 4*c*. $v = 200$.

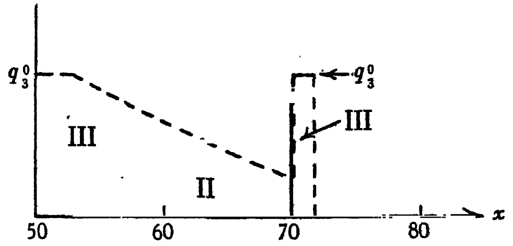


FIGURE 4*d*. $v = 600$.

FIGURE 4*a*–*d*. Development of a band containing one solute with a solvent containing a solute which is less adsorbed. (*a*) Band before development. (*b*) Initial development, beginning with the compression of the band (if q_3^0 is large enough). (*c*) Compression completed, development not quite finished. (*d*) Development of band completed; solute III has overtaken band II.

The various conditions which arise during development are shown in figure 4*a–d*. Eventually band III completely envelopes band II, its frontal concentration being c_3^0. Band II continues to develop according to equation (51) and it is of interest to note that the spreading of the band after travelling a given distance x is considerably smaller than it would have been, if pure solvent had been used.

B. *Development of a band of two solutes*

The indices of the two solutes are chosen so that the faster moving solute has the index 1 and the slower one the index 2. As before, the developing solute has the index 3.

Again we have several cases:

Case 1. $f'_{1(0)} < f'_{2(0)} \leqslant f_{3(c^0_3)}/c_3^0$, where the solute III has no influence on the separation.

Case 2. $f'_{1(0)} < f_{3(c^0_3)}/c_3^0 < f'_{2(0)} < f'_{3(0)}$, where the solute III has only little influence on the separation.

Case 3. $f_{3(c^0_3)}/c_3^0 < f'_{1(0)} < f'_{2(0)} < f'_{3(0)}$ and

Case 4. $f'_{3(0)} < f'_{1(0)} < f'_{2(0)}$.

Both cases 3 and 4 will be discussed in detail.

Case 3

Only the case will be dealt with, where solute III is present in sufficiently high concentration to result in a development corresponding to that in figure 3*b*, as this appears to be of more practical importance.

During the development the band passes through the stages shown in figure 5*a–d*. The original band (figure 5*a*) is compressed during the development to concentrations which are determined by the concentration of the developing solute. At the same time the process of separation starts, so that a moderately complicated system (figure 5*b*) results, which persists until the original concentrations c_1^0, c_2^0 and c_{1y} have disappeared (figure 5*c*). Thereafter the separation continues until it is complete (figure 5*d*).

Apart from the initial compression development, it will be noticed that conditions here are similar to those considered by Wilson, as the concentrations of the bands remain constant during the whole process of separation.

The original concentrations c_1^0 and c_2^0 are in this case of no influence on the separation, as this depends only on the value of $c_2^{0\prime}$ which itself depends on c_3^0. This also concerns the final concentrations c_{2o} and c_{1z}, which, as the bands must be able to move with the same velocity as the forward edge of solute III, adjust themselves to concentrations which are determined by

$$\Delta x/\Delta v = c_3^0/f_{3(c^0_3)} = c_{2o}/f_{2(c_{2o})} = c_{1z}/f_{1(c_{1z})}, \tag{52}$$

or, in the case of the Langmuir equation,

$$\Delta x/\Delta v = (1+b_3c_3^0)/a_3 = (1+b_2c_{2o})/a_2 = (1+b_1c_{1z})/a_1. \tag{52a}$$

If the difference between the adsorption factors of the two solutes to be separated is only small, then—with negligible error—

$$c_1^{0\prime}/f_{1(c^{0\prime}_1 c^{0\prime}_2)} = c_3^0/f_{3(c^0_3)} \tag{53}$$

and

$$(1+b_1c_1^{0\prime}+b_2c_2^{0\prime})/a_1 = (1+b_3c_3^0)/a_3. \tag{53a}$$

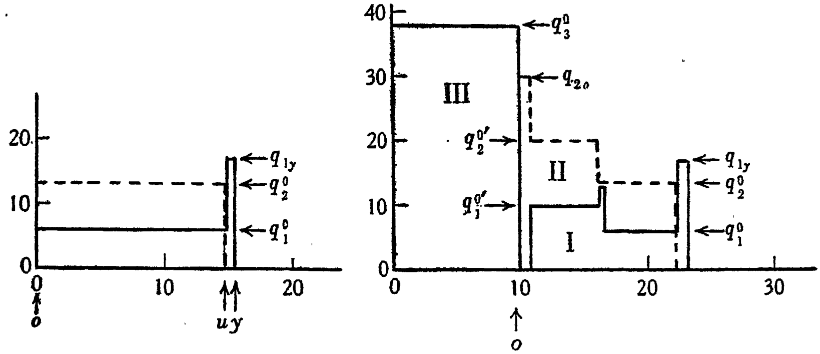


FIGURE 5*a*. $v_0' = 100$, $v = 0$. FIGURE 5*b*. $v = 50$.

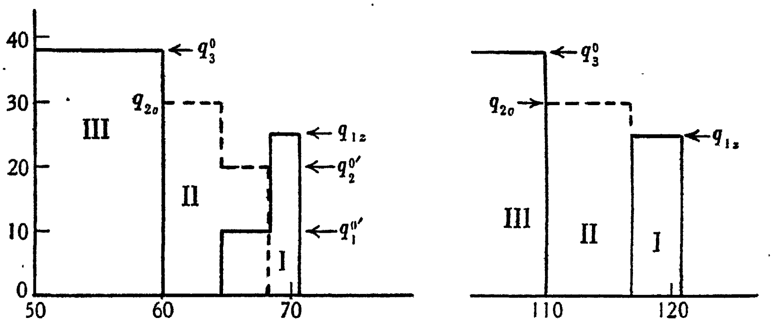


FIGURE 5*c*. $v = 300$. FIGURE 5*d*. $v = 550$.

FIGURE 5*a*–*d*. Development of a band containing two solutes with a solvent containing a stronger adsorbed solute. (*a*) Band before development. (*b*) Compression of band and start of separation. (*c*) Compression completed; separation further advanced. (*d*) Separation completed. Both boundaries are sharp and the band has not spread.

Equation (53*a*), together with the condition

$$c_1^{0\prime}/c_2^{0\prime} = c_1^0/c_2^0 = m_1/m_2, \tag{54}$$

results in

$$c_2^{0\prime} = \frac{\left(b_3c_3^0 - \dfrac{a_3-a_1}{a_1}\right)\dfrac{a_1}{a_3}}{b_1m_1/m_2+b_2}. \tag{55}^*$$

---

* The correct deduction leads to a similar expression in which b_2 is replaced by $b_2\phi$, where $\phi = a_1b_1c_{1z}/(a_2-a_1+a_2b_1c_{1z})$ and where c_{1z} can be replaced by c_3^0 by means of equation (52*a*). If a_1 and a_2 differ only little, ϕ is, under normal working conditions, almost unity.

The volume of solvent V required for complete separation of the bands I and II is (see equation (25*a*))

$$V = \frac{m_2 a_1}{c_2^{0\prime}(a_2-a_1)} = \frac{b_1 m_1 + a_1 b_2 m_2 \phi / a_2}{\frac{(a_2-a_1)}{a_3}\left(b_3 c_3^0 - \frac{a_3-a_1}{a_1}\right)}. \tag{56}$$

The minimum amount of adsorbent required for a complete separation is then (see equation (2))

$$X_0 = V c_1^{0\prime} / f_{1(c^{0\prime}{}_1 c^{0\prime}{}_2)} = V(1+b_1 c_1^{0\prime} + b_2 c_2^{0\prime})/a_1. \tag{57}$$

Again use may be made of the simplifying approximation (53*a*), leading to

$$X_0 = V(1+b_3 c_3^0)/a_3,$$

$$X_0 = \frac{(b_1 m_1/a_1 + b_2 m_2 \phi / a_2)(1+b_3 c_3^0)}{\frac{a_2-a_1}{a_1}\left(b_3 c_3^0 - \frac{a_3-a_1}{a_1}\right)}. \tag{58}$$

A close scrutiny of the meaning of this equation shows that it is almost identical with Wilson's equation

$$X_0 = l_0 \frac{a_2-a_1}{a_1},$$

only that in this case l_0 does not represent the length of the band after formation, but after compression in the course of development by the more strongly adsorbed solute III.

If the process of development is conducted with a sufficiently high concentration of the developing solute, i.e.:

$$1 \ll b_3 c_3^0 \gg (a_3-a_1)/a_1,$$

equation (58) can be simplified into

$$X_0 = (b_1 m_1/a_1 + b_2 m_2/a_2)/\delta. \tag{59}$$

If one compares this with the value obtained for development with pure solvent

$$X_0 = (m_1+m_2)\, b_2/a_1 \delta^2 \quad \text{(see (41}a\text{))},$$

the advantage of developing with a solute is very obvious, if δ is very much smaller than unity. It must, however, be borne in mind that this method leads only to a separation into two adjoining, but not completely isolated bands, because, as has been pointed out, the final concentrations adjust themselves in such a way, that their speed of travel along the chromatographic column is equal. Complete separation into isolated bands can take place only when a method of development resulting in variable concentrations is employed.

Case 4

In this case development takes place similar to that produced by a pure solvent. But, as has been shown before in the case of a single solute, the developed band spreads very much less, especially if solute III is only slightly less adsorbed than the solutes to be separated. In this case the amount of adsorbent required for complete separation may be only little more than that required in case 3, and it provides the possibility for the separation of the solutes I and II into completely isolated bands. As the eventually separated bands I and II will contain substantial amounts of solute III which has to be removed separately afterwards, it is obvious that this procedure only offers advantages for very difficult separations, or, if the substance III can be removed easily by a different process, e.g. by chemical separation or by chromatography with a different adsorbent.

As the general equations, even in the case of the Langmuir isotherm, are fairly cumbersome, a solution will be produced only for the case $(a_2-a_1)/a_1 \ll 1$ and $c_3^0 > c_1^0 + c_2^0$, which assumptions permit considerable simplifications. Equations which have been simplified in this way, will be marked with the letter '*s*'.

The fundamental equations governing the relations between coexisting concentrations of c_1, c_2 and c_3 are again (see equation (2))

$$v/x = f'_{1(c_1c_2c_3)} = f'_{2(c_1c_2c_3)} = f'_{3(c_1c_2c_3)}, \tag{60}$$

where

$$f_{1(c_1c_2c_3)} = a_1c_1/(1+b_1c_1+b_2c_2+b_3c_3) \tag{61}$$

and vice versa. Complete differentiation results in the two equations

$$\left.\begin{aligned} &\left.\begin{aligned} &a_1+a_1b_2c_2+a_1b_3c_3-a_1b_2c_1\,.\,dc_2/dc_1-a_1b_3c_1\,.\,dc_3/dc_1\\ &\quad = a_2+a_2b_1c_1+a_2b_3c_3-a_2b_1c_2\,.\,dc_1/dc_2-a_2b_3c_2\,.\,dc_3/dc_2 \end{aligned}\right\}(62)\\ &\quad = a_3+a_3b_2c_2+a_3b_1c_1-a_3b_2c_3\,.\,dc_2/dc_3-a_3b_1c_3\,.\,dc_1/dc_3. \end{aligned}\right\} \tag{63}$$

The solutions of these differential equations are

$$c_2 = \alpha c_1 + c_{2o} \tag{64}$$

and

$$c_3 = \gamma c_1 + c_{3o}. \tag{65}$$

For indices see figure 6. The values of these four constants—α, γ, c_{2o} and c_{3o}—can be calculated exactly from the two equations obtained by the substitution of c_2 and c_3 into equations (62) and (63) (which process removes c_1 as well):

$$a_1(1+b_2c_{2o}+b_3c_{3o}) = a_2(1+b_3c_{3o}-b_1c_{2o}/\alpha-b_3c_{2o}\gamma/\alpha), \tag{66}$$

$$= a_3(1+b_2c_{2o}-b_1c_{3o}/\gamma-b_2c_{3o}\alpha/\gamma), \tag{67}$$

together with the boundary conditions applied to equations (64) and (65):

$$c_{2d} = \alpha c_{1d} + c_{2o}, \tag{68}$$

$$c_{3d} = \gamma c_{1d} + c_{3o} = 0, \tag{69}$$

where

$$c_{1d}/c_{2d} = m_1/m_2. \tag{70}$$

Further, c_{2o} and c_{3o} are connected by the single-solute equation (50) which for the respective indices assumes the form

$$c_{3o} = c_3^0 - \epsilon c_{2o}, \tag{71}$$

where

$$\epsilon = \frac{a_3b_2c_3^0}{a_2-a_3+a_2b_3c_3^0}.$$

Equations (66)–(71) permit the calculation of c_{2_0} which is required for the facts about the complete separation of the bands I and II (see equations (25) and (25*a*)). The final equation for c_{2_0} after elimination of the other unknown factors is, however, too complicated to give an impression of what is happening, unless numerical examples are used. For this reason a few simplifying assumptions are made, which do not affect the general result. It is assumed that $(a_2-a_1)/a_1 = \delta \ll 1$. This permits us to consider $(c_{1d}+c_{2d})$ as the concentration of a single substance, the value of which could then be obtained from equation (50) for $c_3 = 0$:

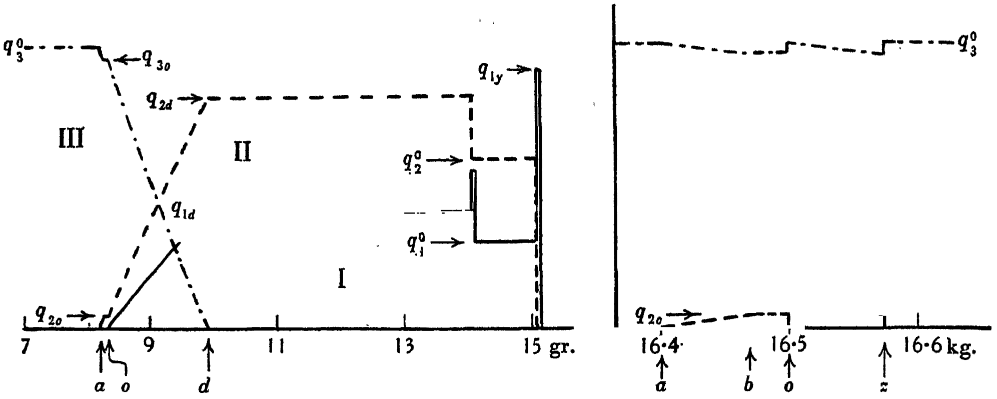


FIGURE 6*a*. $v = 100$ c.c.

FIGURE 6*c*. $v = 200l$.

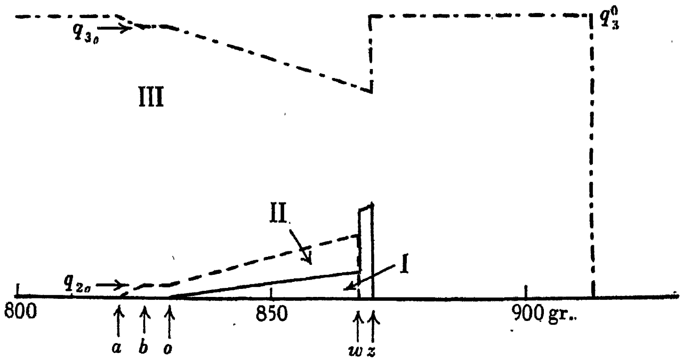


FIGURE 6*b*. $v = 10l$.

FIGURE 6*a–c*. Development of a band containing two solutes with a solvent containing a less adsorbed solute ($v^0 = 100$ c.c., $\delta_{21} = 0{\cdot}01$, $\delta_{23} = 0{\cdot}1$). (*a*) Initial development, preceded by compression of band. Original concentrations have not yet fully disappeared. (*b*) Development complete. Solute III has overtaken the band. Separation well advanced. (*c*) Separation of solutes I and II is complete. Note the relatively large amount of solute III with which the separated solutes are mixed.

$$c_{1d}+c_{2d} = c_3^0/\epsilon. \tag{72s}$$

This additional equation can replace the much more complicated equation (66).

Elimination of c_{1d} and c_{2d} from equations (68)–(72) leads to

$$c_{3o}/\gamma = -m_1 c_3^0/(m_1+m_2)\,\epsilon \tag{73s}$$

and

$$c_{3o}\alpha/\gamma = c_{2o} - m_2 c_3^0/(m_1+m_2)\,\epsilon. \tag{74s}$$

Substitution of c_{3o}, α and γ by means of equations (71), (73), (74) in (67) results in

$$a_1(1+b_2c_{2o}+b_3c_3^0-b_3\epsilon c_{2o}) = a_3\left(1+\frac{b_1m_1c_3^0}{(m_1+m_2)\,\epsilon}+\frac{b_2m_2c_3^0}{(m_1+m_2)\,\epsilon}\right). \tag{75s}$$

The resulting expression for c_{2o} can be very greatly simplified by making the assumption that

$$b_1/a_1 = b_2/a_2 = b_3/a_3 = \beta,$$

which, in the interpretation of the Langmuir equation, means that the surface requirements per mol. of the three solutes are equal. Then

$$c_{2o} = c_3^0\,\frac{b_3[(a_2-a_1)/a_1]\,[1+b_2c_3^0m_2/(m_1+m_2)]}{b_2[(a_2-a_3)/a_3]\,[1+b_3c_3^0]} = c_3^0\,\frac{b_3\delta_{21}\eta}{b_2\delta_{23}}, \tag{76s}$$

where η is likely to have values between 0·5 and 1. The volume V of developing solution required for reaching complete separation of the solutes I and II then follows from equation (25*a*):

$$V = \frac{b_2m_2\delta_{23}}{b_3c_3^0(\delta_{21})^2\eta} \tag{77s}$$

To reach this state requires a minimum quantity of adsorbent

$$X_o = V(1+b_2c_{2o}+b_3c_{3o})^2/a_1$$

$$= \text{approximately } V(1+b_3c_3^0)^2/a_3 \quad \text{(see equation (53}a\text{))}.$$

Thus

$$X_o = \frac{b_2m_2\delta_{23}(1+b_3c_3^0)^2}{a_1\eta(\delta_{21})^2b_3c_3^0}. \tag{78s}$$

To get a minimum value for X_0, the concentration of the developing solute should be chosen so that $b_3c_3^0 = 1$, i.e. half of the adsorbing capacity is filled by substance III. Comparison with the separation conditions obtained without a solute in the developing solvent (see equation (40*b*)) shows that, apart from the factor

$$m_2/(m_1+m_2)\,\eta\, b_3c_3^0,$$

which has little influence unless m_2/m_1 is very small, the decisive improvement in equation (77) is contained in the additional factor $(a_2-a_3)/a_3$. This shows that the use of an extra solute which is less adsorbed than the substances to be separated is only worth while if the adsorption properties of this solute are closely similar to those of the substances to be separated, i.e. if $(a_2-a_3)/a_3 \ll 1$.

The physical reason why the development with another solute gives a faster separation with less adsorbing material, is that the presence of the other solute impedes the developmental spreading of the bands. The fact that, if the concentration of the developing solute is high, the band first undergoes a contraction, is only of a secondary importance; this can be seen from the fact that, in the case of pure solvent development, an initially high concentration of the solutes makes little difference to the complete separation. The decisive fact which emerges clearly from these calculations is that the concentration (c_{2o}) of the band II at the point where separation takes place (x_o) is higher, if the solvent contains a solute which is either more (case 3) or only a little less adsorbed than the substances to be separated.

7. THE CASE OF OTHER ISOTHERMS

It is fully realized that the Langmuir isotherm does not often occur in the case of adsorptions on liquid-solid phase boundaries. This is most probably due to the fact that the adsorbents used do not consist of uniformly adsorbing surfaces, but of a large number of different types of adsorbing centres. Even if every type conforms to a Langmuir equation this would result in an isotherm of the form

$$q_1 = \sum_1^n \frac{a_n c_1}{1+b_n c_1 + b'_n c_2},$$

and, as has been pointed out by Hinshelwood, a number n as small as 2 can already produce a form of curve similar to that of a Freundlich isotherm.

It is apparent that under such conditions where the curvature of the isotherm towards the c-axis is due to saturation of the active centres, the process of chromatographic separation cannot differ greatly from the case of the ideal Langmuir isotherm. The only substantial difference will occur, where the total concentration is very low, i.e. at the trailing end of band II, which will approach zero only asymptotically. However, no exact calculations are possible as the interdependence of the concentrations cannot be determined without solving the differential equation

$$f'_{1(c_1 c_2)} = f'_{2(c_1 c_2)},$$

which in all cases of $n > 1$ leads to differential equations of higher orders. Nor was a solution found possible for any conceivable mixed isotherm apart from the linear Langmuir isotherm. It must therefore be considered as a piece of extremely good fortune, that the only simple isotherm which has a theoretical backing, is capable of a solution and thus gives some insight into the process of chromatographic separation; by way of analogy, this may be useful also for cases where the conditions do not permit of a straightforward solution.

It is, however, of interest to consider the behaviour during a chromatographic separation of substances following an adsorption isotherm which is convex against the c-axis. Such an isotherm would be obtained, for example, under conditions where not only single molecules are adsorbed, but also double molecules of all combinations. Assuming that the adsorbent is far from saturated, this would result in an isotherm of the type

$$q_1 = k_1 c_1 + k_2 c_1^2 + k_3 c_1 c_2. \tag{79}$$

The differential equation (2) cannot be solved for type (79), but, writing this equation

$$q_1 = k_1 c_1(1 + k_2 c_1/k_1 + k_3 c_2/k_1),$$

its form, for $k_2 c_1/k_1 < 1 > k_3 c_2/k_1$, becomes very similar to

$$q_1 = \frac{k_1 c_1}{1 - k_2 c_1/k_1 - k_3 c_2/k_1} = \frac{a_1 c_1}{1 - b_1 c_1 - b_2 c_2}, \tag{80}$$

i.e. equation (3) with negative values of the constants $b_1 b_2$. Here the differential equation (2) can be solved.

As shown by de Vault in his treatment of a single solute, an isotherm convex against the c-axis results in a band with a sharp rear boundary and a diffuse front. In this case the separation condition which can be calculated in a similar way as for a Langmuir equation, leads to a distribution as shown in figure 7.

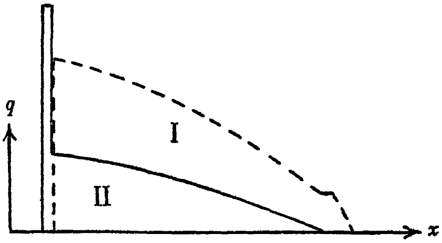


FIGURE 7. Developed band of two solutes with adsorption isotherms which are convex against the c-axis (see equation (80)). The development is almost exactly the reverse of that obtaining in the case of a Langmuir isotherm.

Complete separation is reached, when

$$V + v_0 = \frac{b_1 m_1}{\left(\dfrac{a_2 - a_1}{a_2}\right)^2} \frac{(1+\lambda)}{\lambda} = \text{approx. } b_1(m_1 + m_2) \Big/ \left(\frac{a_2 - a_1}{a_2}\right)^2. \tag{81}$$

The minimum amount of adsorbent required for complete separation is

$$X_0 = \frac{b_1}{a_1} \frac{m_1}{\left(\dfrac{a_2 - a_1}{a_1}\right)^2} \frac{(1+\lambda+\delta)}{\lambda} = \text{approx.} \frac{b_1}{a_1}(m_1 + m_2) \Big/ \left(\frac{a_2 - a_1}{a_1}\right)^2. \tag{82}$$

In equations (81) and (82), λ is defined by equation (8), but with a negative value of b_2.

It is of interest to note that here, too, the amount of adsorbent and solvent required for complete separation depends on $(m_1 + m_2)$ and on $(\delta)^2$ and not, as might be inferred from Wilson's deductions, on δ and m_2 only.

I wish to express my thanks to the Department of Scientific and Industrial Research for permission to publish these investigations, and to Dr A. E. Green, Lecturer in Mathematics at Durham, for solving the differential equation (5) and for discussions concerning the possible use of other adsorption isotherms.

POSTSCRIPT ADDED IN PROOF 14 NOVEMBER 1945

Other general equations for the chromatogram of two solutes formed by normal development

Since this paper went to press, a new fundamental equation of mass conservation has been found for multiple chromatograms (see Glückauf 1945), which makes it possible to give general solutions for most of the equations previously derived for the Langmuir isotherm only, where this has not already been done. These equations still require that the relations between c_1 and c_2 within the mixed band are known, and thus demand the solution of equation 2 by calculation or graphical evaluation.

In the following general equations (marked g) which, for the sake of brevity, are given without deduction and under the same number as the corresponding Langmuir relationships, $F_{1(c_1)}$ and $F_{2(c_2)}$ (see equations 10 and 11) are given the wider meaning

$$F_{1(c_1)} = f_{1(c_1c_2)} \quad \text{and} \quad F_{2(c_2)} = f_{2(c_1c_2)},$$

the mathematical connexion being supplied by equation (2), which requires that c_1 is a function of c_2.

The fundamental equations are:

$$m_{1(o-x)} = xf_{1(c_1c_2)} + x\alpha c_1 - vc_1 \qquad \text{(g 2}a\text{)}$$

$$m_{2(a-x)} = xf_{2(c_1c_2)} + x\alpha c_2 - vc_2 \qquad \text{(g 2}b\text{)}$$

where c_1 and c_2 are the concentrations at the point x, and where

$$x_o < x < x_d \text{ or } x_w.$$

The additional general equations are:

for c_{1y}

$$\frac{c_{1y} - c_1^0}{c_2^0} = \frac{f_{1(c_{1y})} - f_{1(c^0{}_1c^0{}_2)}}{f_{2(c^0{}_1c^0{}_2)}}, \qquad \text{(g 15)}$$

for c_{2w} as function of v (can also be used for v_{du} if $c_{2w} = c_2^0$)

$$v = m_2 \frac{F'_{2(c_{2w})}}{F_{2(c_{2w})} - c_{2w}F'_{2(c_{2w})}}, \qquad \text{(g 28)}$$

for x_w as function of c_{2w}

$$x_w = m_2/(F_{2(c_{2w})} - c_{2w}F'_{2(c_{2w})}), \qquad \text{(g 29)}$$

for c'_{1w} as function of c_{1w} and c_{2w}

$$\frac{c'_{1w} - c_{1w}}{c_{2w}} = \frac{f_{1(c'_{1w})} - F_{1(c_{1w})}}{F_{2(c_{2w})}}, \qquad \text{(g 33)}$$

for the movement of the fully developed frontal band I:

$$x = \frac{v}{f'_{1(c_1)}} + \frac{m_2(f'_{1(c_1)} - F'_{2(c^*{}_{2w})})}{f'_{1(c_1)}(F_{2(c^*{}_{2w})} - c_{2w}F'_{2(c^*{}_{2w})})}, \qquad \text{(g 37)}$$

(c^*_{2w} is a function of c_1; it is defined by equation (g 33) with a value of c_{1w} as following from equation (2) and with $c'_{1w} = c_1$), for the contents of any section of the mixed band:

$$m_{1(p-q)} = v\left[\frac{F_{1(c_1)} - c_1 F'_{1(c_1)}}{F'_{1(c_1)}}\right]_{c_{1p}}^{c_{1q}}, \qquad \text{(g 38)}$$

and vice versa for $m_{2(p-q)}$.

If the pore space is to be taken into account, the corresponding αc has to be added to every function, and α has to be added to every differentiated function.

REFERENCES

Glückauf, E. 1945 *Nature*, **156**, 748.
de Vault, D. 1943 *J. Amer. Chem. Soc.* **65**, 532.
Weiss, J. 1943 *J. Chem. Soc.* p. 297.
Wilson, J. N. 1940 *J. Amer. Chem. Soc.* **62**, 1583.

The plastic deformation of non-cubic metals by heating and cooling

By W. Boas and R. W. K. Honeycombe, *Council for Scientific and Industrial Research, Lubricants and Bearings Section, Melbourne*

(*Communicated by Sir David Rivett, F.R.S.—Received* 1 *September* 1944)

[Plates 1–5]

Several pure metals were subjected to cyclic thermal treatment between 30 and 150° C. Specimens of zinc, cadmium and tin showed signs of plastic deformation which were evident after a small number of cycles, and which became more pronounced as the number of cycles increased. On the other hand, the phenomenon was not observed in lead. Evidence is brought forward to show that this deformation is due to the anisotropy of thermal expansion in the crystals of non-cubic metals. The persistence of the lattice distortions is shown by X-ray photographs, and also by recrystallization subsequent to the cyclic treatment. In certain metals, grain boundary migration is associated with the deformation. The implications of the phenomenon are particularly exemplified in a comparison of certain tin-base and lead-base bearing alloys. Some possible theoretical and practical implications of the phenomenon are discussed.

1. INTRODUCTION

In the course of a comparative study of the properties of tin-base and lead-base bearing alloys, marked differences were observed between the two types of alloy. Small steel cylinders coated on the inside with these bearing alloys were alternately heated and cooled by immersion in oil. After a few cycles between 30 and 150° C, cracking was observed on the surface of the tin-base bearing, but the surface of the lead-base bearing was free from such defects. Moreover, after fifty cycles of heating and cooling, the formerly smooth surface of the tin-base lining had become quite rough. The lead-base lining, however, showed no signs of surface roughening. After two hundred cycles, the general roughening and cracking of the surface of the tin-base bearing had proceeded to a very marked degree. After a similar treatment, no such change was apparent on the surface of the lead-base alloy.

If internal stresses due to the difference in thermal expansion between the steel cylinder and the bearing lining were the cause of these effects, it would be expected that they would be more pronounced the wider the difference in thermal expansion. Although the mechanical properties of tin-base and lead-base bearing alloys are similar (A.S.T.M. 1939), and the difference in thermal expansion between a lead-base alloy and steel is greater than between a tin-base alloy and steel, the surface deformation has been observed only with tin-base alloys. This suggests that the effect was independent of the bimetallic nature of the bearing.

To verify this point, specimens of lead-base and tin-base alloys not bonded to steel were subjected to thermal cycles. Again, the tin-base alloy showed surface deformation whilst the lead-base alloy did not. Figure 1 (plate 1) represents an area on a tin-base bearing alloy which had been subjected to fifty thermal cycles between 30 and 150° C. The rough and deformed state of the originally smooth surface is apparent. These experiments show that the phenomenon is not due to the differential expansion between the steel shell and the bearing metal, but is an inherent property of the alloy, and possibly even of the basis metal. It was therefore decided to investigate the behaviour of pure metals under similar conditions, and specimens of tin and lead were subjected to thermal cycles. After a small number of cycles, tin specimens showed deformation, while no signs of deformation were detected in the lead specimens. The investigation was later extended to cadmium and zinc, and with these metals slip lines were detected even after one thermal cycle.

It was thought that the phenomenon might be due to the anisotropy of thermal expansion, inherent in metals which do not possess a cubic crystal structure. In this paper experimental evidence is advanced to support this explanation.

2. Experimental

(a) *Materials*

The metals used were the purest available. The tin was approximately 99·85 % pure with antimony, arsenic, lead and copper as principal impurities. The lead was of a purity exceeding 99·99 %. The zinc and cadmium used were electrolytically refined and were both 99·97 % pure, with lead and zinc respectively as the main impurities.

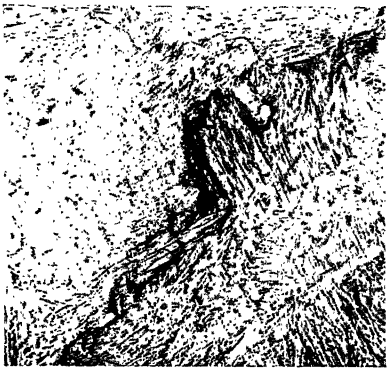
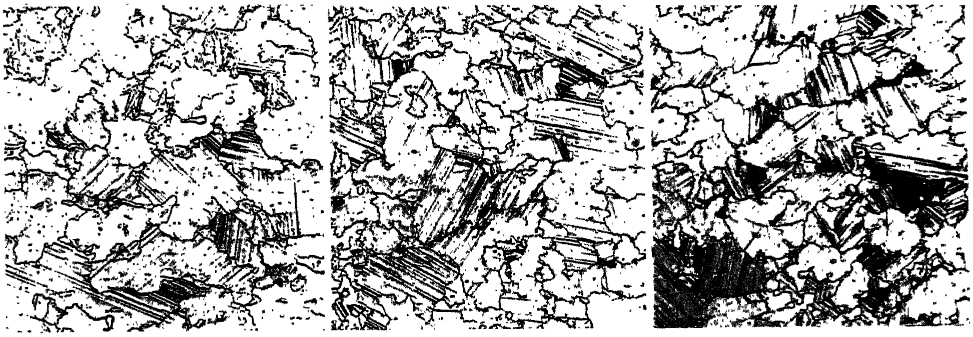
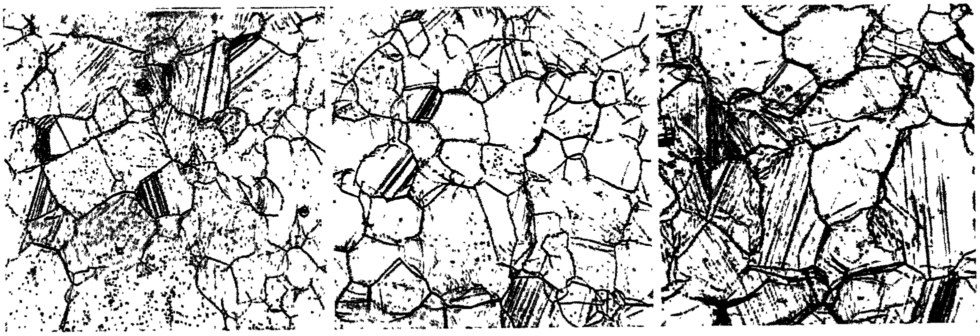


FIGURE 1



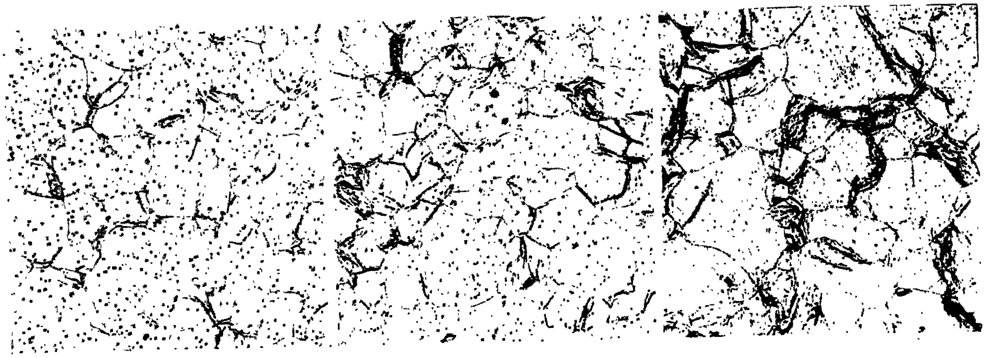
(*a*) (*b*) (*c*)

FIGURE 3



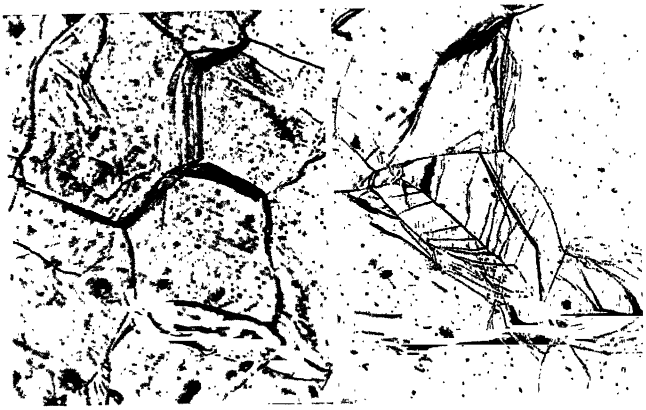
(*a*) (*b*) (*c*)

FIGURE 4



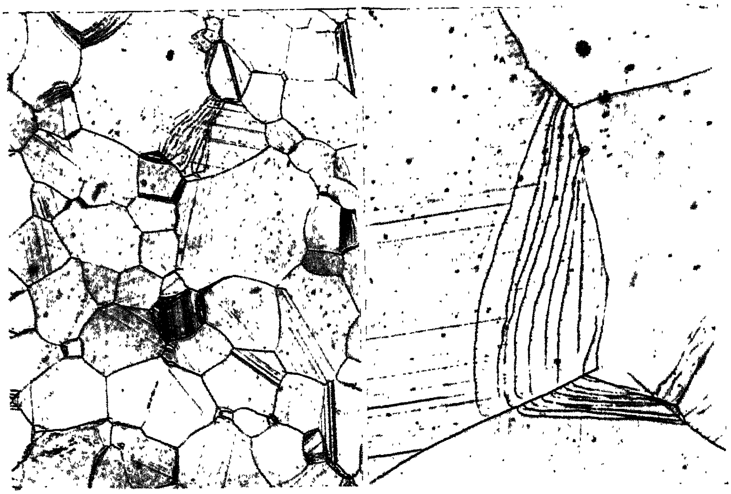
(*a*) (*b*) (*c*)

FIGURE 5



(*a*) (*b*)

FIGURE 6



(*a*) (*b*)

FIGURE 7

Typical compositions of the bearing alloys already referred to were:

	tin-base %	lead-base %
tin	91·9	14·5
antimony	4·0	13·0
copper	4·0	0·5
lead	—	72·0
nickel	0·1	—

(*b*) *Preparation of specimens*

It was essential to obtain the specimens used for the experiments in a state free from any residual stresses which might be produced by mechanical working or polishing. Small ingots of the various metals were cast and rolled to strip approximately 0·20 in. in thickness and 1 in. in width. The specimens were cut from the strip and filed to the shape shown in figure 2, this shape being adopted in order to facilitate handling of the specimen. One surface of each specimen was then rubbed on wet carborundum papers up to 600 grade. The specimens were subsequently annealed for at least 30 min. at a temperature equal to or exceeding the maximum temperature used in the cyclic thermal treatment.

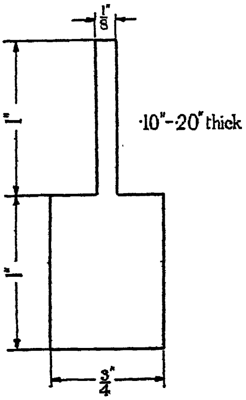


FIGURE 2. Shape of specimens used.

Finally, the specimens were electrolytically polished. The electrolytic method of polishing eliminates the possibility of surface deformation during the preparation of the specimen. Tin specimens were polished by the method of Jacquet (1939), while the cadmium specimens were prepared by a modification of this method. A variation of the technique used by Rodda (1943) was successful in the case of zinc. Lead was at first mechanically polished by the method of Worner & Worner (1940), but later the electrolytic technique of Koch & Stanau (1941) was adopted. Once a specimen was prepared in this manner it was not subjected to further polishing during or after the cyclic thermal treatment.

(c) *Cyclic thermal treatment*

In the earlier experiments, the specimens were alternately immersed in a hot oil bath at 150° C and a cold oil bath at 30° C. A mechanical device was constructed in order to carry out a large number of cycles automatically. In this case the specimens were held in each bath for $2\frac{1}{2}$ min. whilst the removal from one bath to the other occupied 1 min. Originally the specimens were in direct contact with the oil; however, it was later found advisable to enclose them in glass tubes to avoid undue staining. Moreover, the glass containers reduced the cooling and heating rates, thus minimizing the stresses due to a temperature gradient in the specimen. Under these circumstances the duration of the cycle was lengthened. Some experiments were carried out with a still slower rate of heating and cooling, the specimens being heated and cooled in the one oil bath. In this case, the temperature rose to 150° C in 40–60 min., while the cooling to room temperature occupied several hours. The specimens were examined microscopically after a varying number of cycles.

In order to examine the deformation *during* a thermal cycle, specimens were heated and cooled on the stage of a photographic microscope. The specimen was in direct contact with a copper cylinder containing oil which was heated by an immersion heater. The temperature was measured by a thermocouple inserted in a small hole in the side of the specimen. This hole was drilled before the specimen was annealed. The specimen was insulated from the stage of the microscope by a sheet of mica 0·4 mm. thick in which there was a small hole through which the microscopic observations were made.

3. RESULTS

(a) *Description of observed phenomena*

Three types of phenomena were observed in the specimens after cyclic thermal treatment. First, deformation marks appeared in the crystals, the direction of the lines changing at the grain boundaries. The majority of these lines was due to slip although some evidence of twinning was found. This was shown by the facts that slight etching removed most of the lines, and that usually only one set of parallel lines was found in each grain. Twins, on the other hand, are more readily distinguished from the parent grains after etching because of the difference in orientation. Secondly, the smooth surfaces of the specimens became rough and the grain boundaries more marked. Thirdly, migration of boundaries occurred in certain metals, both the old and the new boundaries being visible simultaneously without additional preparation of the specimen. As the number of cycles was increased, the three effects became more pronounced. This is shown for zinc, cadmium and tin in figures 3–5 (plates 1 and 2). The three phenomena did not occur to the same extent in each metal; moreover, they were not all necessarily associated in every case.

Zinc. Slip occurred to a marked degree as is shown in figure 3, but no grain boundary migration was observed. This was accompanied by a roughening of the surface which was readily seen with the naked eye. In addition to this permanent distortion, a reversible roughening of the surface was detected by microscopic observation of most specimens during the cycle, although some specimens did not show this effect.

Cadmium. Slip occurred to approximately the same extent as with zinc. This is shown in figure 4 (plate 1). The surfaces of the cadmium specimens showed permanent roughening which was as pronounced as in zinc. The cyclic treatment brought about accentuation of the grain boundaries, and, in addition, extensive grain boundary migration took place.

Tin. Only slight evidence of slip or twinning was seen (figure 5, plate 2). The main effect appeared to be an accentuation of the grain boundaries, some of which ultimately seemed to become regions of severe distortion as shown in figure 6*a* (plate 2). Extensive grain boundary migration occurred (figure 6*b*), similar to that observed in cadmium.

Lead. No signs of deformation were detected, and no accentuation of the grain boundaries became apparent. In some specimens, certain grains grew at the expense of others, but no grain boundary network was developed. No connexion between this grain adsorption and the cyclic treatment was observed and there was no evidence of recrystallization within the existing grains.

(*b*) *Factors influencing the deformation*

Duration of cycle. Specimens of cadmium which were plunged alternately into hot and cold oil, and specimens which were slowly heated and cooled in the one oil bath, gave similar results. A fast cycle occupied between 5 and 10 min., while a slow cycle took approximately 7 hr. The deformation caused in a cadmium specimen by ten slow cycles is illustrated in figure 7 (plate 2). This photograph also shows more extensive grain boundary migration than is the case in specimens subjected to fast cyclic treatment.

These experiments, together with the fact that no plastic deformation was observed when lead was subjected to a similar treatment, support the view that the deformation is not due to internal stresses caused by non-uniform heating and cooling.

Number of cycles. As already stated, the observed phenomena became more pronounced as the number of cycles was increased. It was also apparent that different grains showed slip to varying degrees. After only one cycle, some grains were already plastically deformed, and as the cyclic treatment proceeded, more and more grains, showed signs of deformation. This progressive increase in the deformation of particular crystals and the onset of deformation in new ones was followed by photographing the same area of a specimen at intervals during the cycles. In figure 8 some photographs taken on a cadmium specimen are reproduced.

Slight indications of deformation may already be observed in specimens as polished electrolytically. This may be seen, for example, in figure 8*a* (plate 3). This might be explained in either of two ways. First, the deformation lines may be remnants of those which are produced on cooling from the annealing temperature, and are not completely removed by the subsequent electrolytic polishing. This explanation is supported by the fact that such specimens do show distinct slip and twinning when microscopically examined after annealing but prior to polishing. This deformation is at times sufficiently pronounced that a rumpling of the surface of the annealed specimens can be detected. For example, the surfaces of cadmium specimens prepared on no. 600 carborundum paper and subsequently annealed at approximately 250° C have a rumpled appearance and signs of plastic deformation can be detected in the grains. The second possible explanation is that the specimens were deformed on cooling from the temperature at which the electrolytic polishing was carried out (usually 45–60° C for cadmium), although the first explanation appears more likely.

In general, not only did an increasing number of grains show deformation lines as the number of cycles was increased, but also the lines became more intense and more numerous in grains deformed in a previous cycle.

Grain size. The effect of variation in grain size was investigated specially in the case of cadmium. It was apparent that deformation occurs independently of grain size. The smallest average linear grain size investigated was approximately 0·05–0·06 mm. (figure 8) and grain sizes up to 100 times this have been examined. Thick grain boundaries observed can be resolved at higher magnifications into a large number of grain boundary impressions which are the result of limited migration.

A significant difference in behaviour was observed with a single crystal of cadmium. The specimen was of the same shape as those previously used. It was subjected to 50 cycles by plunging alternately into oil at 30 and 150° C. Here, one would expect the effect to be absent, and the results showed that this was so. The major part of the surface remained entirely free from any slip lines. Slight indications of deformation were observed at the base of the handle which was polycrystalline, and along one edge where the specimen was subjected to the greatest thermal gradient.

Temperature. Microscopic observations during the cycle revealed that the number of slip lines in a grain increased as the temperature was raised (figure 9, plate 4). Moreover, the intensity of existing slip lines increased as the cycle proceeded and was augmented not only as a result of the rise in temperature but also of the subsequent cooling. Figure 9 shows that the slip lines are at least as pronounced near grain boundaries as they are in the centre of the grain. The deformation is accompanied by grain boundary migration which is also seen in figure 10 and which will be discussed later.

Metal. As the preceding experimental evidence implies, the deformation is very pronounced in the case of zinc and cadmium. It is difficult to judge which of these metals undergoes the greater deformation under similar conditions. On the other hand, tin shows slip to a much lesser degree, but slip and twinning are not necessarily the only outward signs of deformation. The severe distortions near the grain boundaries seem to be the principal indications of deformation in tin. In lead no deformation can be detected.

Orientation of the crystals. It was thought that the effect of relative orientation of the crystals on the deformation could be investigated statistically by comparing the behaviour of specimens in which the crystals were orientated at random with specimens possessing a preferred orientation of the crystals. To this end, two series of zinc specimens were prepared. In the first case, heavy cold rolling in one direction was carried out followed by annealing at approximately 300° C, while the material for the second series was hot rolled in various directions and annealed. After cyclic thermal treatment there was some indication that less plastic deformation occurred in specimens of the first series. However, further work is necessary to confirm this result, as the existence of preferred orientation in these specimens has not yet been established.

(c) *Associated phenomena*

Persistent nature of the lattice distortions. No recrystallization was detected *during* the cyclic treatment. On the other hand, it was possible to bring about recrystallization in specimens of cadmium by annealing at a temperature higher than the maximum temperature of the cycle *after* they had undergone a considerable number of cycles. Specimens, after rolling, were annealed for ½ hr. at 250° C, the resultant grain size being the same in each case. The structure was revealed by macro-etching (figure $10a_1$–e_1, plate 5). They were then given 10, 25, 50, 100 and 200 cycles respectively, reannealed for ½ hr. at 250° C and finally macro-etched. Figure $10a_2$ shows that the deformation produced during 10 cycles is not sufficient to allow recrystallization to occur. Partial recrystallization took place in the specimen previously subjected to 25 cycles (figure $10b_2$) and led to the formation of very large grains. Recrystallization was complete in the specimen which had received 50 cycles (figure $10c_2$), resulting in a somewhat finer grain-size than that of the previous specimen. The specimens subjected to 100 and 200 cycles were completely recrystallized by the annealing treatment (figures $10d_2$ and e_2), and the grain size of both specimens was smaller than in the previous two cases. In other experiments with cadmium, the same general behaviour was observed, but the minimum number of cycles required for recrystallization on annealing varied.

These experiments indicate that some lattice distortions produced by the plastic deformation of the specimen are not removed during the cyclic treatment but remain in the specimen and, therefore, become more extensive as the number of cycles increases.

This opinion is supported by evidence from X-ray diffraction photographs taken on zinc, cadmium and tin at stages during the cyclic treatment. Back reflexion photographs were made using the continuous X-ray spectrum and the characteristic radiation from a copper target. Care was taken to obtain reflexions from the same crystals at different stages during the experiments. As the grain size was relatively large, Laue spots from several crystals were obtained, those produced by the characteristic radiation being particularly intense. In the photographs of the specimens in the as-annealed condition, the majority of the interferences was sharp.

In the three metals after 100 cycles, the spots were blurred, and both the $K\alpha_1$ and $K\alpha_2$ radiations were reflected from the one crystal. This is shown for cadmium in figure 11 (plate 5).

These two effects are identical with those described by Gough & Wood (1936) and indicate that the orientation varies within the one crystal where previously no such lattice distortion was present. The lattice distortions increased with the number of cycles. In certain cases, after a large number of cycles, recovery occurred as is indicated by the splitting up of some of the blurred spots into several smaller ones sharper in outline.

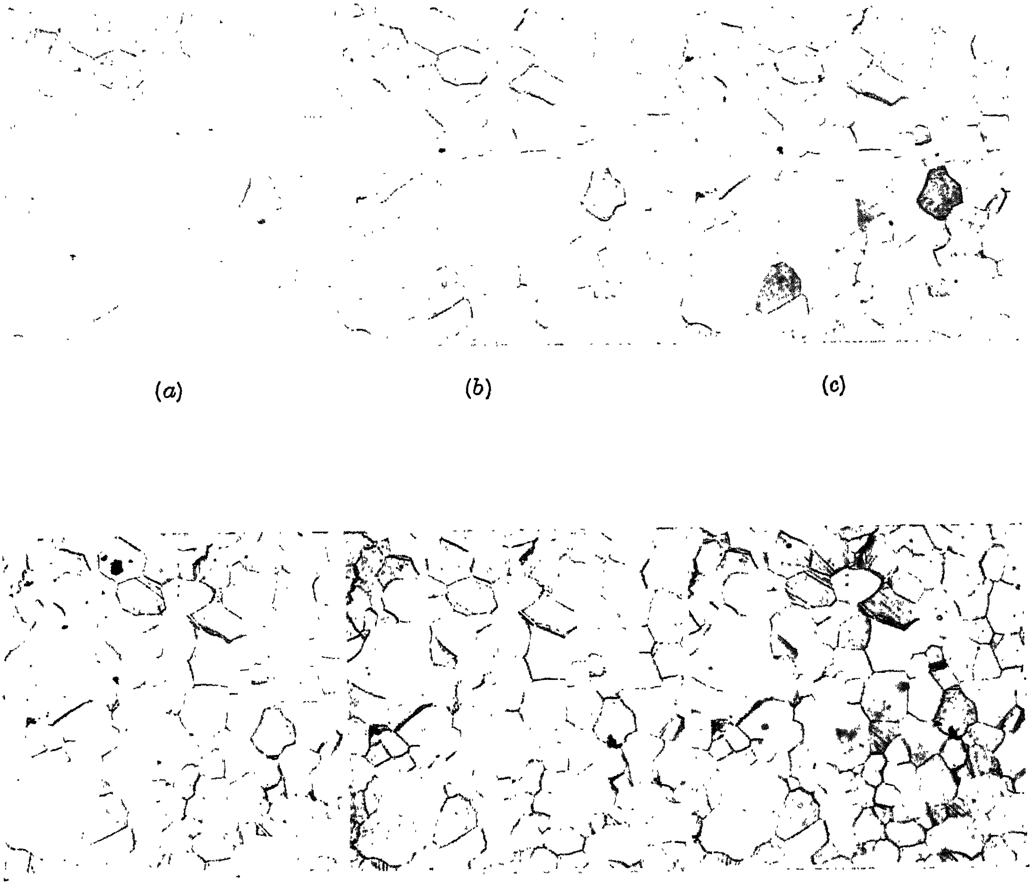
In similar experiments with lead, the sharpness of the Laue spots was not altered by the same cyclic thermal treatment as that applied to the non-cubic metals.

Grain boundary migration. It has already been mentioned that during the cyclic treatment, grain boundary migration occurs in the case of tin and cadmium, but not in zinc. In both tin and cadmium some grain boundaries moved more than others, while some were stationary. In very many instances, a grain boundary moved during each cycle and an impression of the boundary was then formed, ultimately leading to a complicated network of lines. On etching, these networks disappeared leaving only the new grain boundaries. These boundaries coincided with the final impression observed before etching, thus showing that the lines of the network represented the positions of the grain boundaries at various times. The number of grain boundary impressions observed did not always coincide with the number of cycles.

Unlike the deformation, the extent of the grain boundary migration is dependent on the duration of the cycle. In specimens of cadmium subjected to very slow cyclic treatment (7 hr. cycle) the migration was very much more extensive (see figure 7) than in the case of specimens which had received shorter cycles. A closer examination of grain boundary migration was made on cadmium and tin specimens which were microscopically examined during the cycle. Grain boundary migration appeared to occur to a much greater extent during the cooling part of the cycle. This can be seen in the case of cadmium in figure 9. The final state of the specimen after three such cycles is shown in figure 9*g* and at higher magnification in figure 9*h*. These photographs show, moreover, that the direction of the migration may be either into the deformed grain or into the undeformed grain. The slip lines formed after this grain boundary migration extend to the new boundary (figure 7*b*).

4. DISCUSSION AND CONCLUSIONS

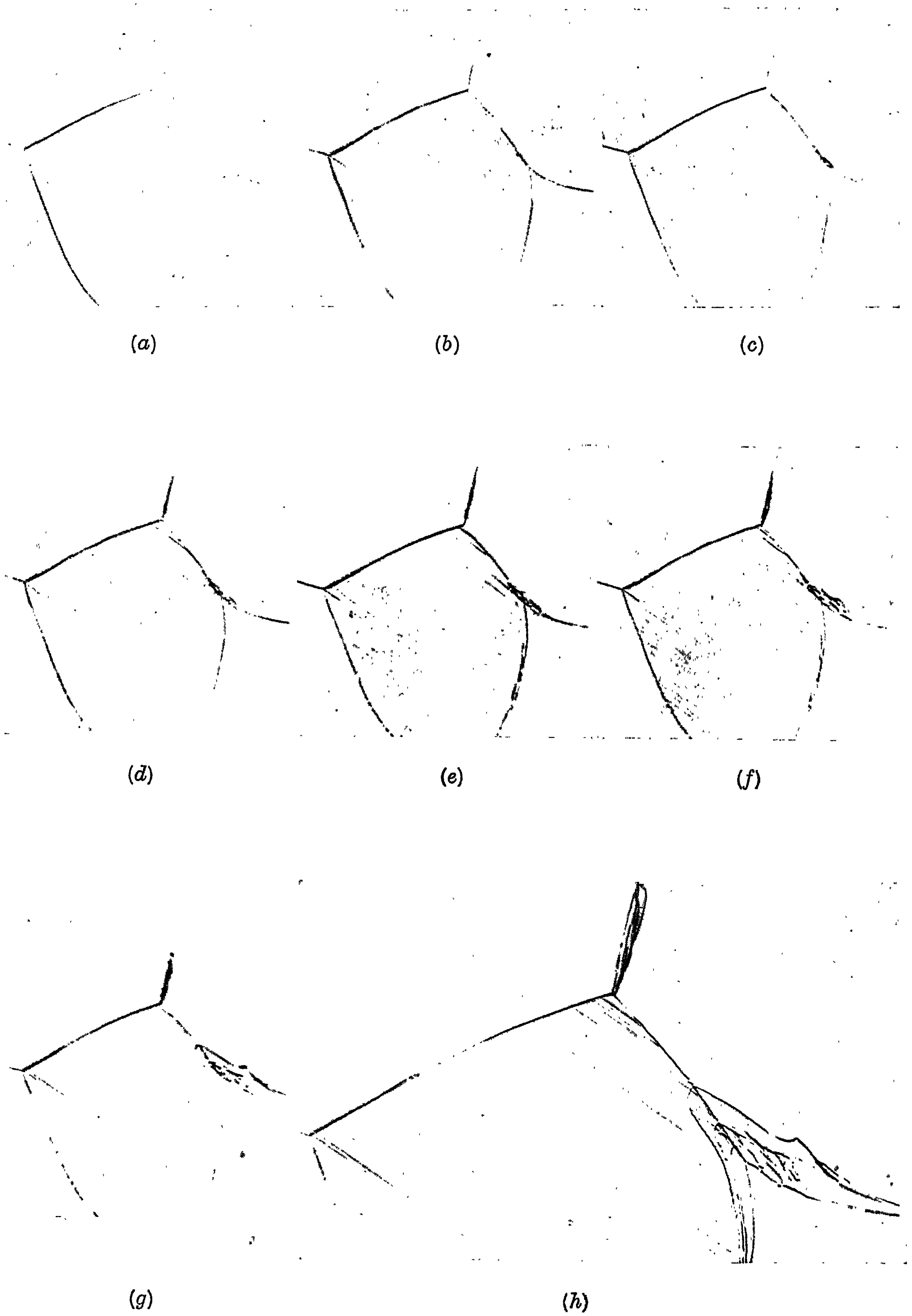
It is considered that the preparation of the specimens prior to cyclic treatment was carried out carefully enough to eliminate residual stresses set up in the metal by previous cold working. These precautions were especially necessary in view of the experience of Rosenhain & Ewen (1912), who observed twins on annealing silver which had been previously mechanically polished. They ascribed the twinning to



(*a*) (*b*) (*c*)

(*d*) (*e*) (*f*)

FIGURE 8

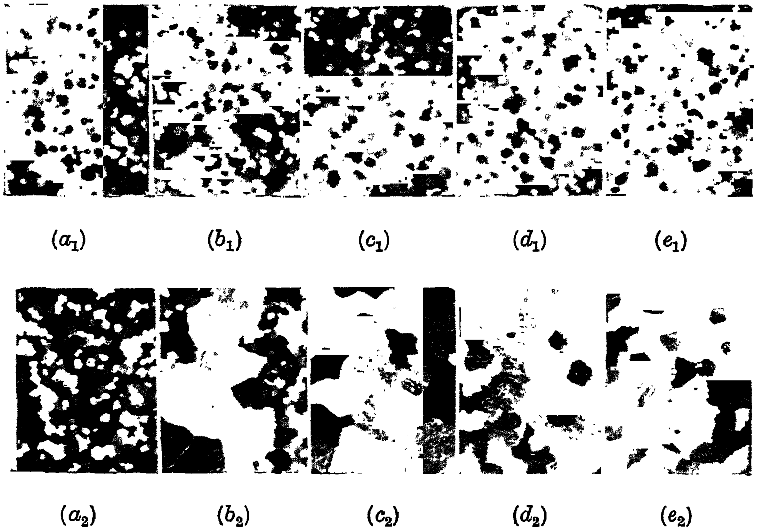


(*a*) (*b*) (*c*)

(*d*) (*e*) (*f*)

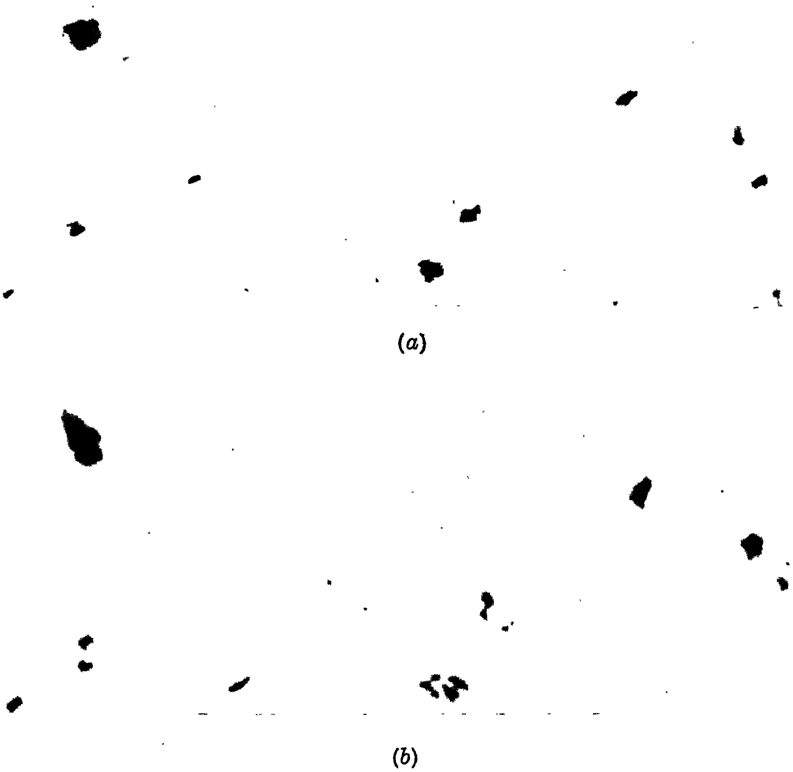
(*g*) (*h*)

FIGURE 9


(a_1) (b_1) (c_1) (d_1) (e_1)

(a_2) (b_2) (c_2) (d_2) (e_2)

FIGURE 10


(a)

(b)

FIGURE 11

the deformation caused by this type of polishing. Those authors did not detect a similar effect with zinc.

It is also unlikely that a temperature gradient in the metal, causing non-uniform expansion in different parts of the specimen, was responsible for the deformation. This is shown in the first instance by the small effect which the duration of the cycle had on the extent of the deformation, and, secondly, by the negative result of the experiments carried out with lead under the same conditions. Since the deformation produced by the thermal cycles was apparently characteristic of the non-cubic metals examined, the possibility of it being due to the anisotropy of thermal expansion became likely.

In metals belonging to the hexagonal and tetragonal crystal systems, the thermal expansion varies with the direction in the crystal; consequently, temperature changes will give rise to stresses where two crystals of different orientations adjoin, and where free expansion is hindered, that is, at the grain boundary. If these stresses are large enough, plastic deformation of the grains will occur. In a cubic metal, on the other hand, no such effect can exist because the thermal expansion is the same in all crystallographic directions.

If it is now assumed that the anisotropy of thermal expansion is the cause of the deformation described above, then it would be expected that:

(1) The cubic metal lead would not show any evidence of plastic deformation after cyclic thermal treatment.

(2) The rates of heating and cooling of the specimens during cyclic treatment would have no effect on the deformation, since the thermal expansion is dependent on the temperature difference and not on the rate at which the temperature changes.

(3) Specimens possessing a strongly preferred orientation would show deformation to a lesser degree than those in which crystals are orientated at random, since, on the average, the change in orientation from crystal to crystal is smaller in the former case than it is in a material with a random distribution of orientations. In the extreme case of a specimen consisting of only one crystal, no deformation would occur.

The first two expectations were borne out in the experiments described above. The third also appears to be supported by the experimental evidence, but this point needs further confirmation. It should be particularly noticed that no effect was observed with the cubic metal lead although its melting point (327·5° C) is very close to that of cadmium (320·5° C), and although it is very easily deformed.

It is therefore considered that the deformation observed in zinc, cadmium and tin is consistent with the anisotropy of thermal expansion in the crystals of these metals.

An estimation of the magnitude of the stresses produced during the heating of a non-cubic metal is found in the appendix. This calculation shows that the anisotropy of thermal expansion, rather than the anisotropy of Young's modulus or the critical shear stress, determines the likelihood of occurrence of plastic deformation. The numerical values show that plastic deformation should occur most readily in zinc, nearly to the same extent in cadmium, and to a lesser extent in tin. The experimental results appear to confirm this expectation. On the other hand, the small number of visible slip lines observed in specimens of tin after cyclic treatment may not give a true indication of the extent of deformation, since a lack of slip lines has been observed in some single crystals of tin which had undergone considerable elongation in tensile tests (Polanyi & Schmid 1925).

It is clear from the experimental evidence that the lattice distortions not only persist, but accumulate as the cyclic treatment proceeds. It might be expected that recovery would take place during each cycle, but apparently the rate of recovery is not sufficiently rapid to eliminate the lattice distortions.

This is not surprising since strain hardening has been observed with single crystals of zinc and cadmium in tensile tests carried out at elevated temperatures (Boas & Schmid 1930). The accumulation of lattice distortions with increasing number of cycles results, on subsequent recrystallization, in a decreasing grain size. This behaviour is analogous to that of metals deformed by external stresses, where the grain size after recrystallization is smaller the greater the extent of the deformation.

It has been previously mentioned that the slip lines continue through the crystals to the grain boundaries. This is in marked contrast to observations on metals deformed in a tensile test. In the latter case, slip lines become less numerous and tend to fade out near the grain boundaries. The emphasis of slip lines near grain boundaries in specimens subjected to cyclic thermal treatment indicates that, in these regions, stresses are unusually heavy, thus supporting the conception of the origin of the deformation outlined above.

In a thermal cycle, as the temperature is raised, the stress between two crystals increases at first proportionately to the temperature until plastic deformation commences. Any greater stress is then largely relieved by plastic deformation; however, a slight increase in stress occurs because of the work hardening. On cooling, the stress decreases again proportional to the decrease in temperature, until zero stress is reached above room temperature. As room temperature is approached, stresses are set up in the opposite direction, which may be large enough to produce further plastic deformation. This is diagrammatically shown in figure 12. During each successive cycle, these variations in stress are repeated. Similar fluctuations of stress occur in specimens which are subjected to fatigue conditions, for example, to cycles of alternating tensile and compressive stresses. In this case, the deformation in the reversed direction occurs by slip along alternate parallel sets of slip planes, and as the number of stress reversals is increased, further slip lines accompanied by lattice distortions are produced. The analogy between the deformation produced during repeated thermal cycles and repeated cycles of externally applied stresses is striking. For this reason, it was thought that 'thermal fatigue' might be a suitable name for the phenomenon (Boas & Honeycombe 1944).

In the experimental results, certain facts were noted which may be of significance to the main phenomenon, although at first they may not seem to have any close connexion with it. The most notable of these effects is the migration of grain boundaries which was observed in both tin and cadmium. Since this grain boundary migration is apparently associated with the cyclic thermal treatment, in so far as further migration occurs during successive cycles, it is possible that the deformation of the metal is a contributing factor. However, in addition to the deformation, the duration of the cycle and the maximum temperature reached seem to play some part. The extent of grain boundary migration increases in the order zinc, cadmium, tin—that is, with decreasing melting point. However, lead, which possesses a melting point very close to that of cadmium, shows little grain boundary migration after cyclic thermal treatment.

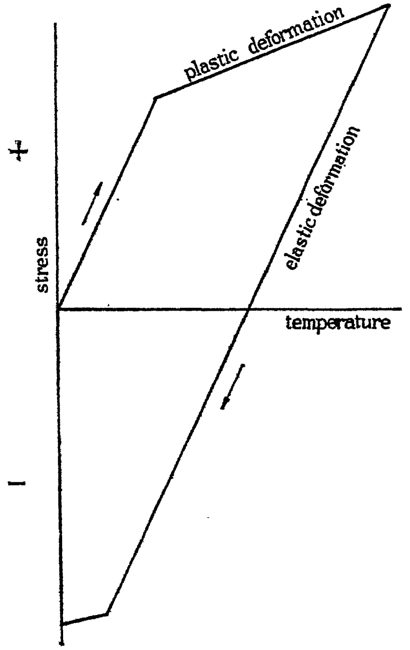


FIGURE 12. Diagram showing the change of stress during the first thermal cycle.

In the course of an investigation on recrystallization and crystal growth, Carpenter & Elam (1920) obtained very similar results with a tin-antimony alloy containing 1·5 % antimony. Those authors used this alloy because of its 'very peculiar property', namely, that 'the position of the new boundary is marked by a line which is really a difference of level, almost as if the specimen had been etched'. They found that these boundary markings were produced independently of whether the specimen was heated in air, hydrogen or in vacuo, and that they were formed only when the annealing was interrupted and the specimen was cooled to room temperature. It is possible that the grain boundary migration observed by those authors had the same origin as that described above. Although Carpenter & Elam concluded that crystal growth would not take place unless a metal had been previously deformed, they found that grain boundary movements occurred when an annealed tin-antimony alloy was reheated. This supports the suggestion that the deformation caused by cyclic thermal treatment is connected with grain boundary migration.

As yet, no reasons have been advanced why the new grain boundaries become visible on cooling. Preferential volatilization at the grain boundaries is unlikely to be the cause, since the excess of the vapour pressure at the grain boundaries over that within the grains is proportional to the surface tension, which is of the same order of magnitude for zinc, cadmium, tin and lead. Neither can the effect be attributed to etching by hot oil, since it was equally observable in air, and, as Carpenter & Elam have shown, in vacuo and hydrogen. It is therefore thought that the impression indicating a difference of level is caused by the stresses set up during cyclic thermal treatment.

Finally, some implications of the phenomenon might be mentioned. It appears that it is difficult to obtain these metals, particularly zinc and cadmium, in a completely strain-free condition at room temperature. Although all stresses due to previous cold work may be eliminated by annealing at a sufficiently high temperature, new stresses due to the anisotropy of thermal expansion will be set up even if the rate of cooling from the annealing temperature is very slow. Similar stresses may occur in these metals on cooling after casting. Preliminary experiments indicate that this is the case. Although no failure in pure metals was observed in the above experiments, it might be brought about after a sufficiently large number of cycles.

The position is somewhat different in the case of alloys containing more than one phase. Here, differences in thermal expansion may occur not only from grain to grain but also from phase to phase, and stresses may be set up even if all the phases present are cubic. On the other hand, the relative volumes taken up by each phase may be of importance as seems to be the case in bearing alloys in which the crystal structure of the matrix is the determining factor. Many bearings are raised to a temperature exceeding 100° C under normal running conditions. Clearly, over a period of time, the bearings will receive cyclic thermal treatment, the severity of which depends on the number of times the bearing is allowed to cool after use. This may ultimately contribute to the failure of certain bearing alloys in service. In this respect, lead-base bearing alloys were found to be superior to tin-base alloys.

It would appear that metals which possess a high degree of anisotropy of thermal expansion cannot be obtained in a strain-free condition at room temperature by casting or annealing. The observations are of some interest in the general theory of the strength of materials and the physical properties of metals. The stresses produced during thermal cycles may be of significance when they are superimposed on externally applied stresses. For example, in the case of fatigue, the stresses in the metal may be raised in this way from the safe to the unsafe range resulting in unexpected failure. It is conceivable that these stresses play a role in the behaviour of metals under creep conditions where the deformation at the grain boundaries is significant.

This work forms part of a general investigation carried out by the Lubricants and Bearings Section of the Council for Scientific and Industrial Research. We wish to express our thanks to Dr F. P. Bowden for his interest and encouragement in the work. We also wish to thank Professor J. Neill Greenwood for helpful discussions and the use of the equipment in the Rosenhain Memorial Laboratory, and to thank Professor E. J. Hartung for laboratory facilities in the Chemistry School of the University of Melbourne.

APPENDIX

Estimation of the order of magnitude of the stresses set up on heating and cooling of non-cubic metals

The basic idea of the calculation of stresses due to the anisotropy of thermal expansion is that two neighbouring crystals possessing different coefficients of thermal expansion must still conform to the same boundary after a change in temperature. This is possible only if the crystal with the larger coefficient of thermal expansion is compressed, and that with the smaller expansion is extended. An exact calculation is difficult, for the stresses depend on the shape of the crystals. Furthermore, each crystal has many neighbours, usually of different orientations, and the total stress is obtained by superimposing the various stress systems. A crude way of estimating the magnitude of the stresses is to consider a linear boundary element between two crystals. A somewhat similar estimation has been carried out by Laszlo (1943).

When the temperature of such a specimen is raised by ΔT the length l of the boundary element would be increased by $\alpha_1 l \Delta T$ or $\alpha_2 l \Delta T$ respectively if the crystals 1 and 2 (with the coefficients of thermal expansion α_1 and α_2 respectively) could expand independently of each other. The strains set up are such that the system has a minimum potential energy, and this is the case when the compressive stress in one crystal is equal to the tensile stress in the other crystal. If f_1 is the fraction of the deformation taken up by crystal 1 (and f_2 the fraction taken by crystal 2) and E_1 and E_2 respectively the Young's moduli of the two crystals, it follows that $f_1 E_1 = f_2 E_2$ and the stress

$$\sigma = (\alpha_1 - \alpha_2) \Delta T \frac{E_1 E_2}{E_1 + E_2}.$$

It can be seen that the difference in thermal expansion between the two crystals is the main factor determining the stress.

In hexagonal and tetragonal crystals, the coefficient of thermal expansion depends on the orientation which, in this case, is solely defined by the angle ϕ between the principal axis and the direction under consideration, i.e. the grain boundary. If

$\alpha_\parallel$ and $\alpha_\perp$ are the coefficients of thermal expansion parallel and perpendicular to the principal axis and $\gamma = \cos\phi$, α in any direction ϕ is given by

$$\alpha = \alpha_\perp + (\alpha_\parallel - \alpha_\perp)\gamma^2,$$

so that

$$\sigma = (\alpha_\parallel - \alpha_\perp)\,\Delta T(\gamma_1^2 - \gamma_2^2)\frac{E_1 E_2}{E_1 + E_2}.$$

It is obvious that the stress disappears not only if both crystals have the same orientation whatever the direction of the boundary, but also if the boundary bisects the angle between the two principal axes. Maximum stress occurs if the boundary is parallel to or perpendicular to the principal axis of one crystal, and the other principal axis is nearly perpendicular to it.

The stress divided by the critical shear S_0 may be taken as a measure of the tendency for slip to occur. The decrease in the critical shear stress with increasing temperature is not very marked and is similar for the different metals. Thus, as a first approximation, the value of S_0 determined for room temperature may be used.

For the purpose of comparing the likelihood of occurrence of slip in the various metals, the values of the stresses set up are calculated for crystals of the same orientation relative to the grain boundaries and for the same temperature range. In this case, ΔT and the factor $(\gamma_1^2 - \gamma_2^2)$ which involves the relative orientation, are identical for the metals compared. In the factor containing the Young's moduli, the orientation is not eliminated. This is seen for hexagonal crystals, for which $E_1 E_2 / E_1 + E_2$ becomes

$$\frac{1}{2s_{11} + (-2s_{11} + 2s_{13} + s_{44})(\gamma_1^2 + \gamma_2^2) + (s_{11} + s_{33} - 2s_{13} - s_{44})(\gamma_1^4 + \gamma_2^4)},$$

with the elastic coefficients s_{ik} in the Voigt notation.

Since the variation of this factor with the orientation does not greatly influence the magnitude of the stress, comparative values for the various metals will be calculated assuming that the grain boundary is parallel to the principal axis in one crystal and perpendicular to that axis in the other crystal. Table 1 contains the values for the elastic, plastic and thermal properties of zinc, cadmium and tin* from which values indicating the relative tendency for slip to occur have been determined. It is seen that this tendency increases in the order tin, cadmium, zinc.

TABLE 1

metal	$\alpha_\parallel - \alpha_\perp$	$E_\parallel$ kg./mm.^{2}	$E_\perp$ kg./mm.^{2}	S_0 kg./mm.^{2}	$\dfrac{\alpha_\parallel - \alpha_\perp}{S_0(1/E_\parallel + 1/E_\perp)}$
zinc	$49{\cdot}8 \times 10^{-6}$	12,130	3,560	0·094	14·6
cadmium	$31{\cdot}2 \times 10^{-6}$	8,300	2,880	0·058	11·6
tin	$15{\cdot}0 \times 10^{-6}$	8,640	2,680	0·133	2·3

---

* Data tabulated by Schmid, E. & Boas, W., *Kristallplastizität* (1935), Berlin: J. Springer.

REFERENCES

A.S.T.M. Standards 1939 Part I, B 23–26, p. 679.
Boas, W. & Honeycombe, R. W. K. 1944 *Nature*, **153**, 494.
Boas, W. & Schmid, E. 1930 *Z. Phys.* **61**, 767–781.
Carpenter, H. C. H. & Elam, C. F. 1920 *J. Inst. Metals*, **24**, 83–131.
Gough, H. J. & Wood, W. A. 1936 *Proc. Roy. Soc.* A, **154**, 510–539.
Jacquet, P. 1939 *International Tin Research and Development Council, Technical Publication*, no. 90.
Koch, L. & Stanau, H. 1941 *Metallwirtschaft*, **20**, 752–755.
Laszlo, F. 1943 *J. Iron Steel Inst.* **147**, 173–199.
Polanyi, M. & Schmid, E. 1925 *Z. Phys.* **32**, 684–712.
Rodda, J. L. 1943 *Mining and Metallurgy*, **2**, 323.
Rosenhain, W. & Ewen, D. 1912 *J. Inst. Metals*, **8**, 171.
Worner, H. K. & Worner, H. W. 1940 *J. Inst. Metals*, **66**, 45–51.

DESCRIPTION OF PLATES

FIGURE 1. Area on a tin-base bearing alloy after 50 cycles between 30 and 150° C. × 50 (7/10 original size).

FIGURE 3. Typical areas on a zinc specimen after various numbers of fast thermal cycles between 30 and 150° C, (*a*) after 50 cycles, (*b*) after 100 cycles, (*c*) after 200 cycles. × 100 (7/10 original size).

FIGURE 4. Typical areas on a cadmium specimen after various numbers of fast thermal cycles between 30 and 150° C, (*a*) after 50 cycles, (*b*) after 100 cycles, (*c*) after 200 cycles. × 100 (7/10 original size).

FIGURE 5. Typical areas on a tin specimen after various numbers of fast thermal cycles between 30 and 150° C, (*a*) after 50 cycles, (*b*) after 100 cycles, (*c*) after 200 cycles. × 100 (7/10 original size).

FIGURE 6. Distortion and grain boundary migration in tin specimens, (*a*) after 100 cycles, × 500; (*b*) after 100 cycles, × 100 (7/10 original size).

FIGURE 7. Slip lines and grain boundary migration in a cadmium specimen after 10 slow cycles, (*a*) × 100; (*b*) × 500 (7/10 original size).

FIGURE 8. An area on a cadmium specimen after various numbers of fast thermal cycles. (*a*) as polished, (*b*) after 1 cycle, (*c*) after 3 cycles, (*d*) after 5 cycles, (*e*) after 10 cycles, (*f*) after 15 cycles. × 100 (7/10 original size).

FIGURE 9. Area on a cadmium specimen at different stages during the first three thermal cycles, (*a*) as polished, room temperature, (*b*) first cycle (150° C), (*c*) room temperature after first cycle, (*d*) second cycle (115° C), cooling down, (*e*) room temperature after second cycle, (*f*) third cycle (150° C), (*g*) room temperature after third cycle, (*h*) same as *g*—higher magnification. (*a*)–(*g*) × 64; (*h*) × 140 (9/10 original size).

FIGURE 10. Macrostructures of cadmium specimens. a_1–e_1 as prepared for cyclic treatment. a_2–e_2 the same specimens, annealed for ½ hr. at 250° C after (a_2) 10, (b_2) 25, (c_2) 50, (d_2) 100 and (e_2) 200 cycles. × natural size.

FIGURE 11. Enlarged portions of X-ray back reflexion diagrams from a cadmium specimen showing the same spots, (*a*) before cyclic thermal treatment, (*b*) after 100 cycles. Copper target.

A contribution to the theory of the Foucault test

By E. H. Linfoot, *University of Bristol*

(*Communicated by C. R. Burch, F.R.S.—Received* 17 *February* 1945)

A diffraction theory of the Foucault test is developed which covers the practically important case of a circular mirror with errors of figure of arbitrary form, tested with a knife-edge whose lateral adjustment is variable. The theory is first applied to discuss the changes in appearance of a true mirror as the knife-edge is advanced across the axis. Next, corresponding formulae are obtained for the case of a mirror possessing a small amount of primary spherical aberration, tested near mean focus. Lastly, the theory is applied to obtain an estimate of the sensitiveness of the test, first as a means of determining focal position, and secondly as a means of detecting primary spherical aberration.

1. Introduction

Although more than eighty years have passed since Foucault published the first account of his knife-edge test, it is still by far the most widely used method of testing astronomical mirrors and other high-quality optical systems of large aperture during the process of figuring. The test is too well known to need a description here; it is only necessary to note that it is of unsurpassed simplicity and delicacy when the stigmatism of a pencil is the property to be tested. For then the interpretation of the test is so easy as to be in a sense intuitive; the shadow contrasts, interpreted as if they arose from the illumination of a nearly flat disk by light coming in at grazing incidence from the side opposite to the knife-edge, depict to the eye the deviations of the wave-surface from a sphere centred where the knife-edge meets the axis. And its sensitiveness is such that errors in the wave-front far too small to cause any observable change in the image at best focus can be clearly seen and accurately located.

In his original description of the test, Foucault (1859) gave a purely ray-theoretic explanation of its properties. Such an explanation serves well enough when large errors are present. Certain unexplained effects are seen, but they can be disregarded without impairing the practical usefulness of the test. As the errors are reduced, however, these effects become more disturbing, and one of them—a brilliant line of light round the rim of the disk—often causes trouble to the inexperienced mirror-maker through being taken (as on ray theory it must be taken) to indicate the presence of a steep, narrow turned edge. Dark fringes appear on the brightly lit areas of the disk seen under the test, and if the knife-edge is kept at a fixed setting there is no satisfactory way of deciding whether or not these indicate a local reversal of the error slope, such as might occur from rapid zonal error on the wave-front.* Fortunately, their characteristic behaviour as the knife-edge is moved from side to side makes it easy to distinguish these diffraction fringes, which evanesce and reform in different positions as the knife-edge is moved, from 'true knife-edge shadows', which change their form and intensity much more slowly. The usual practice is to advance the knife-edge until no diffraction fringes remain on the zone under examination and then to use the ray-theoretic interpretation.* It is far from evident, however, that the shape of the zone profile is thereby correctly inferred from its variations in brightness when the errors are comparable with the wave-length of light.

---

* Such rapid errors are often met with during the zonal figuring of heavily aspheric plates.

An adequate diffraction theory of the Foucault test is therefore desirable on practical as well as on theoretical grounds. Moreover, it is a necessary basis for any estimate of the theoretical limit to the sensitiveness of the test.

The first approach to such a theory was made by Rayleigh (1917). Using a simplified two-dimensional model, he explained, qualitatively at least, the bright rim shown by an error-free mirror under the test and the rather similar bright line which appears at a step-discontinuity on the mirror surface.

Zernike (1934), in a remarkable paper† on the knife-edge test and the phase-contrast test, developed a theory covering errors of arbitrary form, but under the restriction that the optical retardations should be small compared with the wave-length of light. Gascoigne (1945), taking the Zernike theory as his starting point, obtained expressions for the knife-edge intensities, valid for errors of arbitrary form and amount, in terms of definite integrals (Hilbert transforms) of a very simple type. He applied these to a discussion of the fringes which appear on the 'bright slopes' when turned edge or zonal errors amount to several wave-lengths. Though derived, like Zernike's, for the simplified two-dimensional model,‡ his formulae yield results of great practical value connecting the depths of errors of certain commonly occurring types with the number of fringes observed.

In the present paper the three-dimensional diffraction theory of the Foucault test is developed in a form which allows the prediction of the appearance under test of a mirror possessing smooth errors of figure of arbitrary shape, and is applied to discuss some questions of importance to the practical optician.

In §§ 2·1, 2·2 the mathematical formulation of the problem is set up, and it is shown that for the present purpose the effect of the finite aperture of the viewing system can be disregarded. § 2·3 shows that, when this is done, the three-dimensional theory can be made to depend in a very simple way on the two-dimensional theory and establishes, as three-dimensional analogue of a result of Gascoigne, a general formula for the intensities seen under the test on a mirror of circular edge contour with arbitrary errors of figure.

In § 3 this formula is applied to a discussion of the changes in the appearance of a true circular mirror under the test as the lateral setting of the knife-edge is varied.

---

* 'I pushes the waves off the mirror with the knife-edge and then I tests with the rays alone.'

† Equations (18) and (20) of this paper are rendered incorrect by an error in sign in equation (14).

‡ The second part of Zernike's paper, in which he discusses the phase-contrast test, is three-dimensional.

In § 4 the general case is again considered and series expressions are obtained which give the intensities seen under test, on the mirror disk and in the halo, in terms of coefficients which describe the errors on the mirror. These expressions are used in § 5 to estimate the theoretical limits of sensitivity of the test, first as a means of determining focal position, and secondly as a means of detecting primary spherical aberration.

2. GENERAL THEORY OF THE FOUCAULT TEST

2·1. All the cases where an aberration-free pencil is being null-tested can be covered by supposing that the wave-fronts originate at the surface of a spherical mirror which is being knife-edge tested at its centre of curvature. The wave leaving the surface of the mirror at the point (x, y) is described by an electric wave-displacement

$$E(x,y)\exp\left[\frac{2\pi i}{\lambda}ct\right],$$

the real part of which measures the electric intensity at (x, y) at the time t. Here λ denotes the wave-length, c the velocity of light. To express the fact that a wave-surface of constant intensity and phase is leaving the surface of a circular mirror M of radius 1, we set

$$\begin{aligned} E(x,y) &= 1 \quad (x^2+y^2 \leqq 1),\\ &= 0 \quad (x^2+y^2>1). \end{aligned} \tag{2·11}$$

Of course, such a discontinuous distribution of electric displacement values cannot be strictly realized.

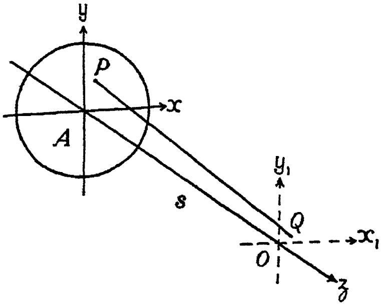


FIGURE 1

The more general case where the pencil under test is not fully stigmatic can be covered by supposing that spherical wave-fronts, originating from a pinhole near its approximate centre of curvature, are reflected from the surface of a nearly spherical mirror M. The results of the knife-edge test are then interpreted in terms of errors of figure of the mirror M. We suppose that these errors of figure amount to several reflexion fringes,[*] but that the error slopes on the mirror, besides being free from discontinuities, are not so steep as to spread out the visible image to more than a moderate multiple, say 5 or 10, of the size of the Airy disk. Then it makes no appreciable difference if one supposes that the wave is leaving a true spherical surface M_0, lying everywhere within a few wave-lengths of the surface M, and that at the point (x, y, z) on M_0 the complex displacement is

---

[*] Since errors of figure on the mirror are transferred with a factor 2 to the reflected wave-fronts, it is convenient to express them in 'reflexion fringes', a term borrowed from interferometry; a reflexion fringe is half a wave-length.

$$E(x,y) = |E(x,y)| \exp\left[\frac{2\pi i}{\lambda}\phi(x,y)\right]. \tag{2·12}$$

$|E(x,y)|$ is then the amplitude at the point (x,y) on M_0 and $\phi(x,y)$ the phase there; the variation in $\phi(x,y)$ expresses the distortion of the wave-fronts.* The fact that a circular mirror of diameter 2 is illuminated with uniform intensity may be expressed by setting

$$\begin{aligned} |E(x,y)| &= 1 \quad (x^2+y^2 \leqq 1), \\ &= 0 \quad (x^2+y^2 > 1), \end{aligned} \tag{2·13}$$

that is, by defining $E(x, y)$ as zero over the part of M_0 outside the circle C which corresponds to the boundary of the mirror M. $\operatorname{Exp}\left[\frac{2\pi i}{\lambda}\phi(x,y)\right]$ is the phase function which describes the errors of M.

After leaving the mirror, the light comes to a more or less imperfect focus in the neighbourhood of O, the centre of curvature of M_0. See figure 1. By an application of Huyghen's principle† the wave-displacement is calculated in the 'intermediate image surface', namely, the sphere S, centred at A, which passes through the focal point O. The knife-edge will later lie in the tangent plane at O to this sphere. It is supposed throughout that the pencil under test is of small angular aperture, that is to say, the focal distance $AO = s$ is supposed large compared with the diameter of M. Then the distance between a point $Q = (x_1, y_1)$ near O in the surface S and any point $P = (x, y)$ on the part of M_0 enclosed by C is approximately given by the equation‡

$$PQ = s - \frac{xx_1 + yy_1}{s}. \tag{2·14}$$

The contribution to the wave-displacement at Q from the element $dx\,dy$ of area situated at P on M_0 is thus§ a constant multiple of

$$E(x,y) \exp\left[\frac{2\pi i}{\lambda}ct\right] \exp\left[\frac{2\pi i}{\lambda}\frac{xx_1+yy_1}{s}\right] dx\,dy,$$

---

* The approximation here consists in the assumption that the amplitude is unchanged in passing from the point (x, y) on M to the point (x, y) on M_0.

† A discussion based on Maxwell's equations meets with formidable mathematical difficulties; see Poincaré (1892, 1897).

‡ The position of P is given by its x, y space co-ordinates; its z co-ordinate is a first-order small quantity, since s is large. The position of Q in the surface S is similarly defined by the space co-ordinates x_1, y_1 (see figure 1). The object of the device, due to Michelson (1905), of the spherical intermediate image surface S is to secure that (2·14) shall remain correct, to within a small fraction of a wave-length, over an area of the intermediate image surface which is large compared with the Airy disk of M.

§ To replace $d\sigma$ by $dx\,dy$ is here equivalent to varying the amplitude by a factor which runs from 1 at the centre of the mirror to cos $1/s$ at its edge. For a mirror working at $f/10$ this corresponds to a variation in the intensity of illumination of less than ¼ %.

and the total wave-displacement at Q is

$$\frac{1}{2\pi}\iint_M E(x,y)\exp\left[\frac{2\pi i}{\lambda}\frac{xx_1+yy_1}{s}\right]dx\,dy$$

times a constant multiple of $\exp\left[\frac{2\pi i}{\lambda}ct\right]$. It is convenient to omit factors of this form and to use the term 'complex displacements' for the expressions which remain when they are removed. Thus a complex displacement is indeterminate to the extent of a constant normalizing factor.

On writing

$$\frac{2\pi x_1}{\lambda s} = u, \quad \frac{2\pi y_1}{\lambda s} = v \tag{2·15}$$

and dropping a constant factor, one obtains for the complex displacement at Q the expression

$$\begin{aligned} W(u,v) &= \frac{1}{2\pi}\iint_{x^2+y^2\leqq 1} E(x,y)\,e^{iux+ivy}\,dx\,dy \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(x,y)\,e^{iux+ivy}\,dx\,dy, \end{aligned} \tag{2·16}$$

since $E(x,y)$ is defined as zero for $x^2+y^2>1$.

It is easy to verify that, in the case (2·11) of a true mirror, (2·16) reduces to the well-known expression

$$J_1[\sqrt{(u^2+v^2)}]/\sqrt{(u^2+v^2)};$$

the intensity in the surface S is then measured by the quantity

$$|W(u,v)|^2 = \frac{J_1^2[\sqrt{(u^2+v^2)}]}{u^2+v^2}. \tag{2·17}$$

This agrees with the usual formula for the Airy disk and rings in the *plane* through O perpendicular to the axis, which will be called the (x_1, y_1)-plane. As is well known, it follows from (2·17) that when λs is small* nearly all the light crosses S in a small region surrounding the point O.

2·2. To carry out the knife-edge test, the part of the (x_1, y_1)-plane defined by $x_1 \leqq 0$ is covered by an opaque screen. Immediately behind O, a viewing system L receives the light and images M_0 on to the final image surface M_0' (see figure 2). L may be the observer's eye-lens and M_0' his retina, but for testing a mirror of long focal ratio a viewing telescope immediately behind the knife-edge is desirable; L then consists of this telescope together with the observer's eye- or camera-lens. In any case, the effect of L is to give equality of optical paths between all pairs of conjugate points P, P' on M_0, M_0'.† In particular, the optical path-lengths $(PQ \ldots P')$

---

* Its value is 0·0004 for a mirror working at $f/10$ and for $\lambda = 2\times10^{-5}$.

† See p. 692 of Zernike (1934) for a proof of this statement.

and $(PO \dots P')$ are equal for every position of Q. It follows that the optical path difference

$$(Q \dots P') - (O \dots P') = -PQ + PO = \frac{xx_1 + yy_1}{s}, \tag{2·21}$$

if it be supposed, as it may be for sufficiently large s, that the approximation (2·14) is usable throughout the part of the intermediate image surface S from which light is passed by L.* A system of co-ordinate numbers (x', y') in the surface M_0' is defined by assigning to the image P' of the point $P = (x, y, z)$ in M_0 the co-ordinate numbers $x' = x, y' = y$. Thus the intensity at P' is the same thing as the 'intensity seen under the test' at P, and the path difference (2·21) can be written $(x'x_1 + y'y_1)/s$.

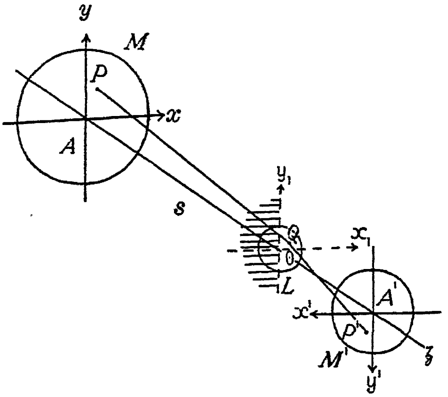


FIGURE 2

The complex displacement $D(x', y')$ at the point (x', y') of M_0' is then obtained by integrating (2·16), multiplied by

$$\exp\left[-\frac{2\pi i}{\lambda}\frac{x'x_1 + y'y_1}{s}\right] = \exp\left[-iux' - ivy'\right]$$

over that part Γ of the intermediate image surface S from which light is passed by the knife-edge and received by the lens-system L. In saying this we make the usual assumption, which cannot be strictly true, that the effect of a screen is merely to stop those parts of the wave which impinge upon it, without influencing the neighbouring parts.† Then

$$D(x', y') = \frac{1}{2\pi}\iint_\Gamma e^{-iux' - ivy'}\, W(u, v) \sec\left(\frac{\lambda}{2\pi}\sqrt{(u^2 + v^2)}\right) du\, dv; \tag{2·22}$$

---

* It can be shown that the errors of L do not affect the general run of the intensity distribution in the final image surface; they merely result in a loss of sharpness comparable (though not identical) with that which would be observed in the image of a self-luminous object lying in the surface of M. But as an adequate discussion of this point would lead too far afield, the problem is here idealized by supposing that the system L is error-free.

† The error involved in this assumption is small provided that the Airy disk is large compared with the wave-length λ, i.e. provided that the angular aperture of the pencil under test is small.

the factor $\sec\left(\frac{\lambda}{2\pi}\sqrt{(u^2+v^2)}\right)$ allows for the fact that the element of area in the surface S is not $dx_1 dy_1 = \left(\frac{\lambda s}{2\pi}\right)^2 du\,dv$ but this quantity multiplied by the secant of the angle between the (x_1, y_1)-plane and the tangent plane to S at the point (u, v). The constant factor $(\lambda s/2\pi)^2$ is omitted in accordance with the present convention.

To replace the factor $\sec\left(\frac{\lambda}{2\pi}\sqrt{(u^2+v^2)}\right)$ by 1 is equivalent to introducing a small variation in the amplitude of the complex displacement $W(u, v)$, which in the case of a 12 in. circular mirror M working at $f/10$ and a viewing telescope of 1 in. aperture amounts to less than one part in 100,000 at the edge of Γ. In this case R, the value of $\sqrt{(u^2+v^2)}$ corresponding to the semicircular edge of Γ, is approximately 1300. The effect of dropping the factor $\sec\left(\frac{\lambda}{2\pi}\sqrt{(u^2+v^2)}\right)$ is therefore negligible and (2·22) can be replaced by the equation

$$D(x', y') = \frac{1}{2\pi}\iint_{u \geqq 0,\, u^2+v^2 \leqq R^2} e^{-iux'-ivy'}\, W(u, v)\, du\,dv. \tag{2·23}$$

The next step is to replace R by ∞ in (2·23); this corresponds to disregarding the effects of the finite aperture of the viewing system. Now it is not true that the replacing of R by ∞ in (2·23) causes only a small change in the value of $D(x', y')$ over the whole mirror, since (as may be shown without great difficulty) the effect of the change is to make $D(x', y')$ tend in general to infinity as the point (x', y') approaches the boundary circle C' corresponding to the edge of the mirror. But if this circle be covered by a narrow annulus, then, provided R is sufficiently large, the change in $D(x', y')$ on replacing R by ∞ is small throughout the remainder of the surface M_0'.* It follows that, except in the immediate neighbourhood of C', the value of $D(x', y')$ is given with sufficient accuracy for the purposes of visual knife-edge testing by the equation

$$D(x', y') = \frac{1}{2\pi}\int_0^\infty du \int_{-\infty}^\infty e^{-iux'-ivy'}\, W(u, v)\, dv,$$

provided that the test is carried out with a viewing system of sufficiently large aperture. It remains to examine whether the apertures ordinarily used in practice are sufficiently large in this sense. For this purpose, an iris diaphragm was fitted to the front of the viewing telescope. In the actual experiment the mirror, of 5 cm. diameter and 275 cm. radius of curvature, was illuminated by an artificial star (pinhole) of diameter 0·006 mm.,† placed close to its centre of curvature. The

---

* An analytical proof of this result will not be included here. Its truth follows from the fact that if, in the surface M_0, the wave-function $E(x, y)$ is given the value

$$\frac{1}{2\pi}\int_0^\infty du \int_{-\infty}^\infty e^{-iux-ivy}\, W(u, v)\, dv$$

and the knife-edge removed, then the complex displacement in M_0' is again given by (2·23).

† I.e. about one-tenth of the diameter, 134λ, of the Airy disk.

aperture of the iris could be varied from 3 to 16 mm.; its plane was 2 cm. behind that of the knife-edge. An aperture of 3 mm. in the plane of the knife-edge would correspond to the value $R = 43$, one of 16 mm. to the value $R = 230$. The magnification of the viewing telescope was 6.

The experiment was tried both with the true mirror and with small errors introduced into the pencil under test by placing pieces of annealed plate glass in front of the mirror. These errors appeared under the knife-edge test as flutings on the apparent surface of the mirror. They increased the overall size of the visible image at best focus to about three times its value for the true mirror. Several different settings of the knife-edge were used in each trial.

When the effect of varying the aperture of the iris was examined, it was found that, over the whole range of variation, no change could be detected in the appearance of the interior of the disk or of the halo. The only visible changes, as the iris was closed down, were a softening in the outline of the brilliant outer rim of the disk, a decrease in its maximum intensity, and the appearance of a fine dark fringe, just inside the outer rim and separating it from the interior of the disk. At minimum aperture, traces of a second dark fringe could be detected inside the first. In the case of the true mirror, a change in the appearance of the rim could first be distinctly seen when the iris aperture was reduced to about 6 mm. (90 Airy disks); this corresponds to the value $R = 95$ and to an exit-pupil diameter of 1 mm. In the case of a pencil with small errors, the aperture usually needed to be reduced a little below this value before the effect could be seen with certainty.

We conclude that, except in the immediate vicinity of the edge contour C', the intensity $I(x', y')$ in the final image surface is given, with accuracy more than sufficient for the purposes of visual knife-edge testing, by the equation

$$I(x', y') = 4\pi^2 \, |D(x', y')|^2, \tag{2·24}$$

where

$$D(x', y') = \frac{1}{2\pi} \int_0^\infty du \int_{-\infty}^\infty e^{-iux' - ivy'} \, W(u, v) \, dv. \tag{2·25}$$

2·3. We suppose that $E(x, y)$ is a differentiable function of x and y at all points (x, y) inside C. Outside C it is differentiable since its value is everywhere zero. On substituting from (2·16) into (2·25) and inverting the order of integration,

$$D(x', y') = \frac{1}{4\pi^2} \int_0^\infty e^{-iux'} \, du \int_{-\infty}^\infty e^{ixu} \, dx \int_{-\infty}^\infty e^{-ivy'} \, dv \int_{-\infty}^\infty e^{ivy} \, E(x, y) \, dy.$$

By Fourier's integral formula, the inner repeated integral

$$\int_{-\infty}^\infty e^{-ivy'} \, dv \int_{-\infty}^\infty e^{ivy} \, E(x, y) \, dy = 2\pi E(x, y'),$$

except at points (x, y') for which $x^2 + y'^2 = 1$. Thus

$$2\pi D(x', y') = \int_0^\infty e^{-iux'} \, du \int_{-\infty}^\infty e^{ixu} \, E(x, y') \, dx. \tag{2·31}$$

It follows from (2·31) that the intensity seen under test at the point P on the mirror depends only on the values of $E(x,y)$ along the horizontal line through P, and is correctly given by applying two-dimensional diffraction theory along this line. On making this application,

$$2\pi D(x',y') = \pi E(x',y') + i\int_{-\infty}^{\infty} \frac{E(t,y')}{t-x'}\,dt \tag{2·32}$$

for all (x',y') not on C', the integral being interpreted as a Cauchy principal value when (x',y') is inside C'. (2·32) can be derived very easily from (2·31) by appealing to a result of Gascoigne (1945, equation 2·8). Since Gascoigne's argument involves the expansion of the wave-function in Legendre polynomials and some use of Bessel functions, the following alternative proof of (2·32) may be of interest.

In (2·31)

$$\int_0^{\infty} e^{-iux'}\,du \int_{-\infty}^{\infty} e^{ixu}\,E(x,y')\,dx = \lim_{U\to\infty} \int_0^{U} e^{-iux'} \int_{-B}^{B} e^{ixu} E(x,y')\,dx,$$

provided B is chosen greater than 1. The expression under the limit sign

$$\begin{aligned}\int_0^{U} e^{-iux'}\,du \int_{-B}^{B} e^{ixu}\,E(x,y')\,dx &= \int_{-B}^{B} E(x,y')\,dx \int_0^{U} e^{-iu(x'-x)}\,du \\ &= -i\int_{-B}^{B} \frac{E(x,y')}{x'-x}\left(1-e^{-iU(x'-x)}\right)dx \\ &= i\int_{-B-x'}^{B-x'} \frac{E(x'+t,y')}{t}\left(1-e^{iUt}\right)dt \\ &= i\int_{-2B}^{2B} \frac{E(x'+t,y')}{t}\left(1-e^{iUt}\right)dt,\end{aligned}$$

since $E(x,y)=0$ whenever $|x|>1$,

$$= i\int_0^{2B} [E(x'+t,y')-E(x'-t,y')]\,(1-\cos Ut)\frac{dt}{t} + \int_0^{2B} [E(x'+t,y')+E(x'-t,y')]\frac{\sin Ut}{t}\,dt. \tag{2·33}$$

Now let $U\to\infty$. By an argument familiar in the classical convergence theory of Fourier series, the second term tends to the limit $\pi E(x',y')$ at all points (x',y') at which $E(x,y)$ is a differentiable function of x, that is to say, at all points (x',y') not on C'.

In the first term of (2·33)

$$\frac{E(x'+t,y')-E(x'-t,y')}{t} = 2\frac{\partial}{\partial x'}E(x',y') + O(t)$$

as $t\to 0$. It follows by an application of the Riemann-Lebesgue theorem that

$$\int_0^{2B} \frac{E(x'+t,y')-E(x'-t,y')}{t}\cos Ut\,dt \to 0$$

as $U \to \infty$. Therefore

$$2\pi D(x',y') = \pi E(x',y') + i\int_0^{2B} [E(x'+t,y') - E(x'-t,y')]\frac{dt}{t}$$

$$= \pi E(x',y') + i\int_0^{\infty} [E(x'+t,y') - E(x'-t,y')]\frac{dt}{t}, \qquad (2\cdot34)$$

provided (x',y') does not lie on C'. But

$$\int_\epsilon^\infty [E(x'+t,y') - E(x'-t,y')]\frac{dt}{t} = \left(\int_{-\infty}^{x'-\epsilon} + \int_{x'+\epsilon}^{\infty}\right)\frac{E(t,y')}{t-x'}\,dt$$

for every $\epsilon > 0$; making $\epsilon \to 0$, we obtain from (2·34)

$$2\pi D(x',y') = \pi E(x',y') + i\int_{-\infty}^{\infty} \frac{E(t,y')}{t-x'}\,dt,$$

which is (2·32).

This equation, together with (2·24), gives the intensity distribution seen under the knife-edge test on a mirror whose wave-form is described by the arbitrary function $E(x,y)$, subject only to the restrictions:

(1) the diameter of the mirror subtends only a small angle at the focal point,

(2) the errors of figure of the mirror, though they may amount to many wavelengths, are small compared with the focal distance s,

(3) the errors of slope on the mirror surface are small and, except at points on the boundary C, the function $E(x,y)$ is a continuous function of (x,y), differentiable with respect to x and y.

It will be seen from the proof of (2·32) that its validity is not confined to the circular mirrors of uniform reflecting power which are discussed in the present paper. The equation is valid, in the same sense as above, for mirrors of arbitrary edge contour, including central piercings or irregularly shaped obstructions, and of variable reflecting power, whenever the maximum angular diameter of the pencil is small and the circumstances are such as to justify disregarding the effect of the finite aperture of the viewing system.

Equation (2·32) reduces the problem of computing the knife-edge intensities to the carrying out of a pair of single integrations at each point (x',y'). In the general case of an arbitrarily given wave-function $E(x,y)$ it is practicable, though decidedly tedious, to carry out these integrations by numerical methods. There are some physically important special cases, however, in which the integrals can be evaluated in terms of known functions.

The simplest such case, that of a true mirror tested with the knife-edge centrally set, is obtained by setting

$$E(x,y) = 1 \quad (x^2+y^2 \leq 1),$$

$$= 0 \quad (x^2+y^2 > 1),$$

as in (2·11). (2·32) then yields at once the formulae

$$\begin{aligned} 2\pi D_0(x',y') &= \pi - i\log\left|\frac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right| \quad (x'^2+y'^2<1)\\ &= -i\log\left|\frac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right| \quad (y'^2\leqq 1<x'^2+y'^2)\\ &= 0 \quad (1<y'^2); \end{aligned} \tag{2·35}$$

$$\begin{aligned} I_0(x',y') &= \pi^2+\log^2\left|\frac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right| \quad (x'^2+y'^2<1)\\ &= \log^2\left|\frac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right| \quad (y'^2\leqq 1<x'^2+y'^2)\\ &= 0 \quad (1<y'^2), \end{aligned} \tag{2·36}$$

which predict the intensities on the mirror-disk and in the halo. The halo is therefore confined to the horizontal band $-1<y'<1$ of the surface M_0'. Along a thin strip y' = constant lying in this band, the intensity function can be written

$$\begin{aligned} I_0 &= \pi^2+\log^2\left|\frac{1+X}{1-X}\right| \quad (|X|<1)\\ &= \log^2\left|\frac{1+X}{1-X}\right| \quad (|X|>1), \end{aligned}$$

where $X = x'\sqrt{(1-y'^2)}$ runs between -1 and 1 in that part of the strip which lies inside C'. It follows that the isophotal lines are ellipses with the vertical diameter BB' as a common axis (see figure 3), and that, in the three-dimensional as in the two-dimensional model,* the halo contains one-third of the total illumination.

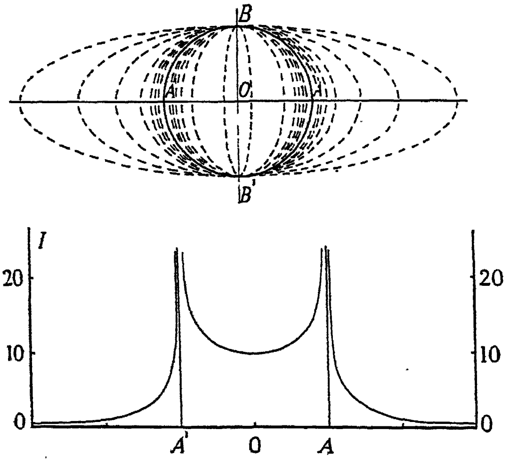


FIGURE 3. Above: Isophotal lines for a true mirror under the Foucault test, with knife-edge central. Below: Intensity distribution along horizontal diameter.

---

* See Zernike (1934), p. 698.

Along the vertical diameter BB' in figure 3, the intensity remains constant; there is therefore no logarithmic infinity at the ends of this diameter. Taken together with the run of the isophotal lines in the interior of the disk, this leads one to expect that in practice the brilliance of the outer rim will not be uniform, but will decrease from a maximum at A, A' to a minimum at B, B'. A practical test confirms this expectation. The effect becomes much more conspicuous when the knife-edge is advanced beyond the central position.*

By applying Babinet's principle, the intensity distribution for a mirror with a pierced central hole can at once be derived from (2·35). The result, for a mirror of inner and outer radii a and b respectively, is

$$I(x',y') = \pi^2\left[\epsilon\left(\frac{r'}{b}\right)-\epsilon\left(\frac{r'}{a}\right)\right]+\left[\epsilon\left(\frac{y'}{b}\right)\log\left|\frac{x'+\sqrt{(b^2-y'^2)}}{x'-\sqrt{(b^2-y'^2)}}\right|-\epsilon\left(\frac{y'}{a}\right)\log\left|\frac{x'+\sqrt{(a^2-y'^2)}}{x'-\sqrt{(a^2-y'^2)}}\right|\right]^2, \tag{2·37}$$

where $\epsilon(t)$ is defined as 1 or 0 according as $|t| \leqq 1$ or $|t| > 1$, and $r' = \sqrt{(x'^2+y'^2)}$. The intensity along the horizontal diameter is shown in figure 4. It is of some interest that its value falls to zero at the centre of the hole when the knife-edge is centrally set.

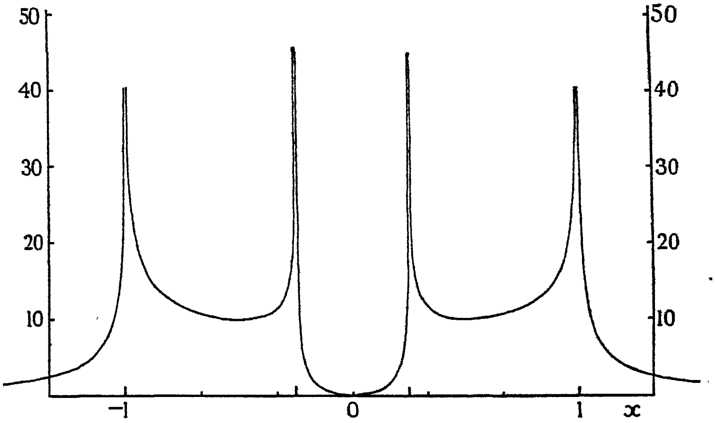


FIGURE 4. Pierced mirror tested with knife-edge central. Intensity along horizontal diameter.

To calculate the effects of small displacements of the knife-edge on the intensities, they may be expressed as 'induced errors' of figure on the mirror. Thus, to determine the effect of advancing the knife-edge across the z-axis through a distance c, $E(x,y)$ is replaced on the right of (2·32) by $e^{ic'x}E(x,y)$ and $D(x',y')$ on its left by $e^{ic'x'}D(x',y')$, where $c' = 2\pi c/\lambda s$. The resulting equation gives the new value of $D(x',y')$. The effect of a change δs in the focal setting can be calculated in a similar way by writing $e^{i\alpha(x^2+y^2)}E(x,y)$ for $E(x,y)$ and $e^{i\alpha(x'^2+y'^2)}D(x',y')$ for $D(x',y')$ in (2·32), where $\alpha = \frac{\pi}{\lambda s^2}\delta s$.

---

* A photograph showing the effect in this case will be found in an interesting paper on edge-diffraction effects by Banerji (1919).

The Reduction Theorem

The following result has now been established:

When a mirror of circular edge contour is tested with a vertical knife-edge, the intensity observed along each horizontal line of the mirror can be predicted to a sufficient approximation by applying two-dimensional diffraction theory along this line. In particular, Gascoigne's equation (2·8), *with* $E(x,0)$ *in place of* $e^{iR(x)}$, *gives with sufficient accuracy the values of the intensities along the horizontal diameter of a circular mirror whose wave-function is* $E(x,y)$.

From this Reduction Theorem, as it may be called, we see that the placing of a horizontal slit-mask over the mirror leaves the intensities sensibly unchanged in the exposed parts of the disk and of the halo, and that Gascoigne's (1945, §§4, 5) computational results can be taken over as they stand into the three-dimensional theory, where they give the knife-edge intensities along the main diameter of a circular mirror suffering from turned edge, zonal error, and hollow centre, in a selected set of special cases.

3. THE TRUE MIRROR

To determine the appearance of a true circular mirror, tested with the knife-edge in the focal plane* but displaced laterally a distance c from the central setting, one must write in (2·32)

$$\begin{aligned} E(x,y) &= e^{ic'x} \quad (x^2+y^2 \leqq 1) \\ &= 0 \qquad (x^2+y^2 > 1), \end{aligned}$$

where $c' = 2\pi c/\lambda s$, and multiply $D(x',y')$ on the left by the phase factor $e^{ic'x'}$. It follows that the complex displacement is given for $|y'| \leqq 1$ by the equations

$$2\pi e^{ic'x'}D(x',y') = \genfrac{}{}{0pt}{}{\pi}{0} e^{ic'x'} + i\int_{-\sqrt{(1-y'^2)}}^{\sqrt{(1-y'^2)}} \frac{e^{ic't}}{t-x'}\,dt,$$

i.e.

$$2\pi D(x',y') = \genfrac{}{}{0pt}{}{\pi}{0} - i\int_{x'-\sqrt{(1-y'^2)}}^{x'+\sqrt{(1-y'^2)}} e^{-ic'u}\frac{du}{u}, \qquad (3·1)$$

according as $x'^2+y'^2 \lesseqgtr 1$, while for $|y'| > 1$, $D(x',y') = 0$.

When $c' = 0$, this reduces to (2·35) above; when $c' \neq 0$, the integral can be evaluated in terms of the functions Si and Ci to give the formulae

$$\begin{aligned} 2\pi D(x',y') = \genfrac{}{}{0pt}{}{\pi}{0} &- \mathrm{Si}\{c'[x'+\sqrt{(1-y'^2)}]\} + \mathrm{Si}\{c'[x'-\sqrt{(1-y'^2)}]\} \\ &- i\,\mathrm{Ci}\{|c'[x'+\sqrt{(1-y'^2)}]|\} + i\,\mathrm{Ci}\{|c'[x'-\sqrt{(1-y'^2)}]|\}, \qquad (3·2) \end{aligned}$$

$$\begin{aligned} I(x',y') = &\left[\genfrac{}{}{0pt}{}{\pi}{0} - \mathrm{Si}\{c'[x'+\sqrt{(1-y'^2)}]\} + \mathrm{Si}\{c'[x'-\sqrt{(1-y'^2)}]\}\right]^2 \\ &+ [\mathrm{Ci}\{|c'[x'+\sqrt{(1-y'^2)}]|\} - \mathrm{Ci}\{|c'[x'-\sqrt{(1-y'^2)}]|\}]^2. \qquad (3·3) \end{aligned}$$

---

* I.e. at the longitudinal setting which produces left-to-right symmetry in the intensities seen under the test.

Some of the consequences of these formulae are of physical interest. When $x'^2+y'^2>1$, i.e. when the point (x', y') is outside the boundary curve C', (3·3) shows that the intensity in the halo is an even function of c', as well as of x' and y'. It is a familiar fact that, as the knife-edge is moved across the Airy disk, the halo brightens to a maximum and then fades out as the mirror disk darkens. The conclusion may therefore be drawn that the halo is brightest when the knife-edge is central. Figure 5 shows the intensity at the point $x' = \frac{11}{10}$, $y' = 0$ in the halo, for different values of c'; when $c' = 3{\cdot}83$ the knife-edge just covers the Airy disk and the first Airy dark ring.

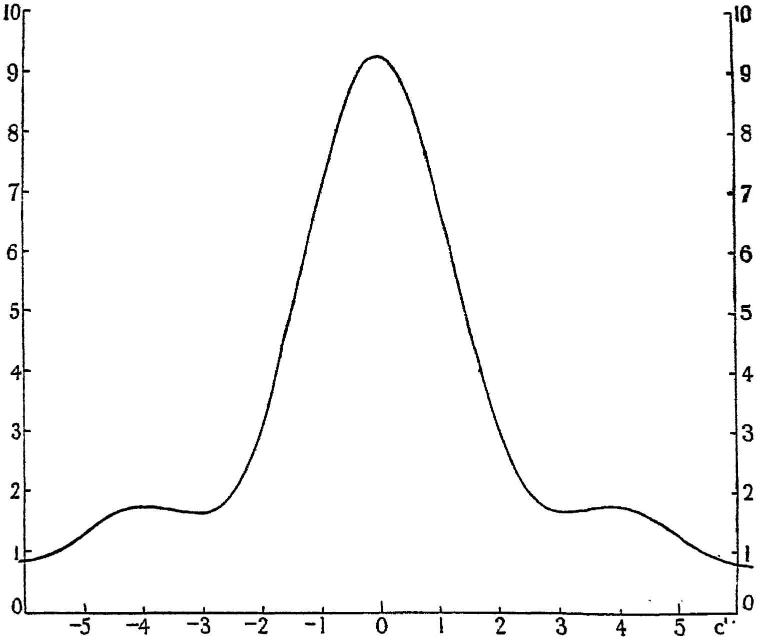


Figure 5. Halo-intensity at $(\frac{11}{10}, 0)$ for different lateral settings of the knife-edge.

Of more practical interest are the changes in the brightness of the disk as the knife-edge is advanced. Banerji (1918) drew attention to the large fluctuations in the observed brightness on the surface of a rectangular mirror which occur as the knife-edge is advanced, and explained them by an application of Rayleigh's theory. Later writers on the practical side of the knife-edge test are almost unanimous in asserting that, as the knife-edge is advanced, a true mirror with circular boundary darkens evenly all over,* and published photographs of the test often show a nearly uniform illumination.† The above discussion shows that, as the knife-edge cuts into the image, undulatory fluctuations of intensity should be seen on the surface of the mirror, and formula (3·3) predicts their appearance when the test is carried out with

---

* Ellison (1939) and Porter (1939), to quote only two examples.

† See, for example, Couder (1937).

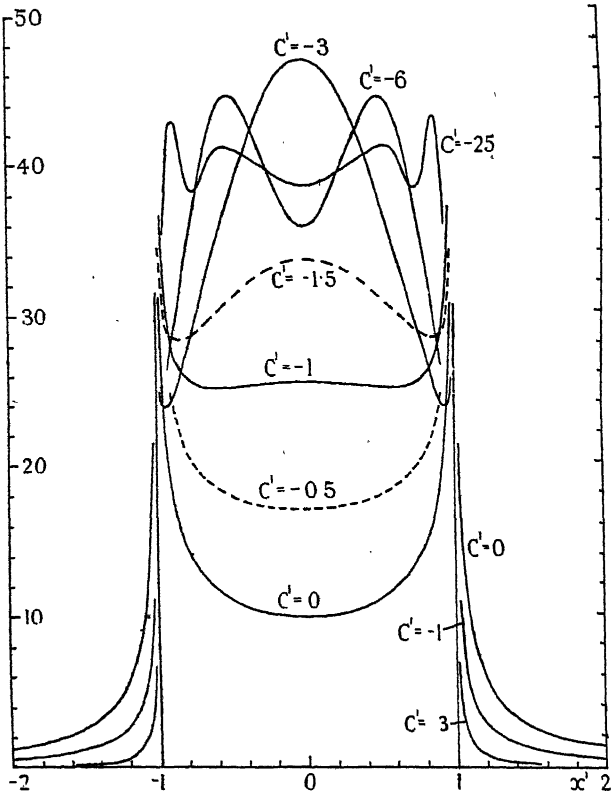


FIGURE 6*a*. True mirror with knife-edge non-central; intensities along horizontal diameter.

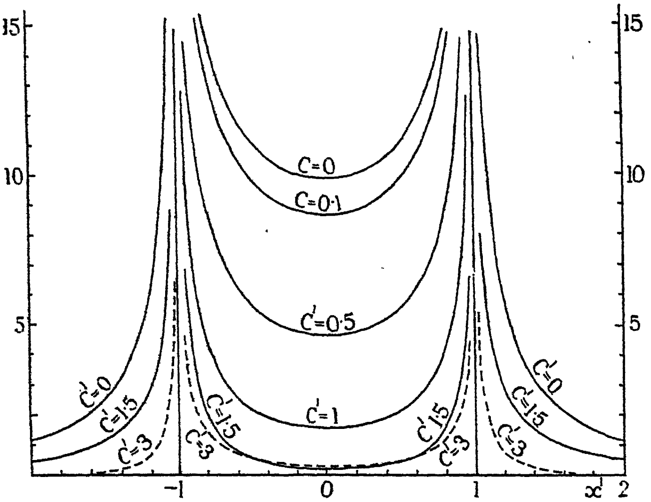


FIGURE 6*b*. True mirror with knife-edge non-central; intensities along horizontal diameter.

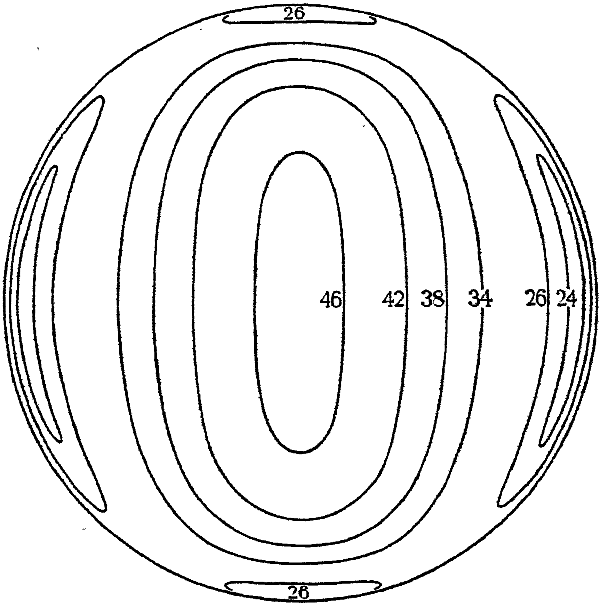


FIGURE 7*a*. True mirror with knife-edge non-central ($c' = -\pi$).

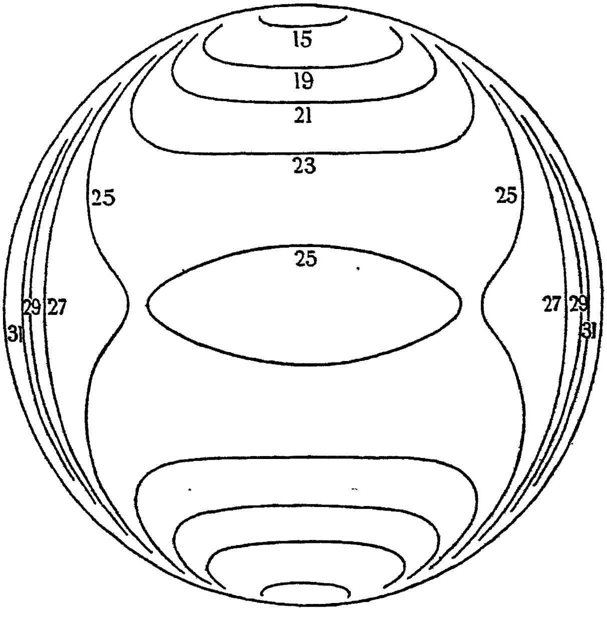


FIGURE 7*b*. True mirror with knife-edge non-central ($c' = -1$).

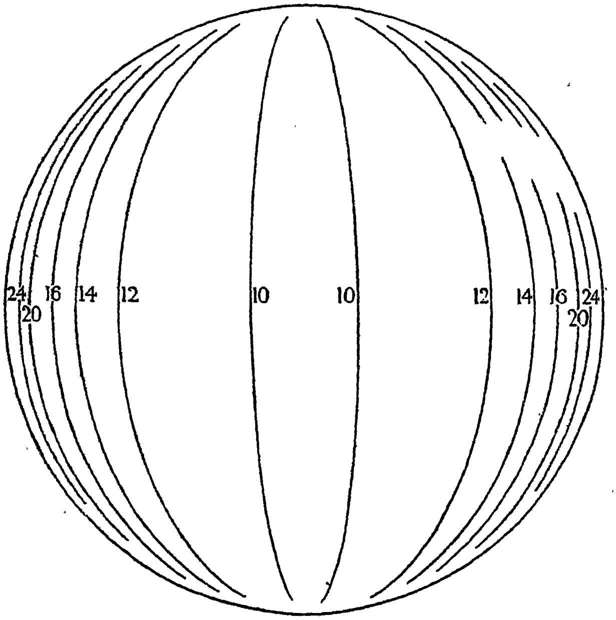


FIGURE 7*c*. True mirror with knife-edge central ($c' = 0$).

a pinhole whose diameter is small compared with that of the Airy disk. Figure 6*a* shows the changes in intensity distribution which take place along the horizontal diameter as the knife-edge advances up to the central position, figure 6*b* those as the knife-edge advances beyond it. It will be seen that, far from darkening evenly all over, the mirror actually brightens in the centre when $c' = -\pi$, i.e. when about one-tenth of the horizontal diameter of the Airy disk is occulted. Figure 7*a* shows the appearance of the disk at this setting. As the knife-edge is advanced, the central intensity maximum flattens out and at $c' = -1$ the disk appears fairly uniformly illuminated (see figure 7*b*). Its appearance when the knife-edge is advanced to the central position is shown in figure 7*c*; the darkening of the central part is then easily perceptible to the eye. The changes described above were observed experimentally without any difficulty on using an $F/55$ spherical mirror, a pinhole of diameter 0·006 mm., and a viewing telescope of aperture 2·5 cm. and magnification 6, placed 2 cm. behind the plane of the knife-edge. It is perhaps relevant that when the eye judges the disk to be half lit, the knife-edge is not central but near to the setting of figure 7*b*, while when the knife-edge is centrally set (by observing the halo intensity or with the help of a microscope) the disk appears considerably less than half lit and the darkening of its centre is easily recognized. In point of fact, as was shown above, the disk contains only one-third as much illumination at this setting as with the knife-edge withdrawn, the remaining one-sixth being sent into the halo.

4. CIRCULAR MIRROR WITH ARBITRARY ERRORS

4·1. To apply (2·32) to the discussion of arbitrary errors on a circular mirror, we suppose that the function $E(x,y)$ can be represented on the mirror by a double power series in x and y, and write

$$\begin{aligned} E(x,y) &= \sum_{m,n\geqq 0} a_{mn}x^m y^n \quad (x^2+y^2\leqq 1) \\ &= 0 \quad (x^2+y^2>1). \end{aligned} \tag{4·11}$$

Then if $-1\leqq y'\leqq 1$

$$\begin{aligned} \int_{-\infty}^{\infty}\frac{E(t,y')}{x'-t}\,dt &= \int_{-\sqrt{(1-y'^2)}}^{\sqrt{(1-y'^2)}}\frac{E(t,y')}{x'-t}\,dt \\ &= \sum_{m,n\geqq 0} a_{mn}y'^n\int_{-Y}^{Y}\frac{t^m\,dt}{x'-t}, \end{aligned}$$

where Y stands for $\sqrt{(1-y'^2)}$. Here

$$\begin{aligned} \int_{-Y}^{Y}\frac{t^m\,dt}{x'-t} &= \int_{x'-Y}^{x'+Y}(x'-u)^m\frac{du}{u} \\ &= \sum_{0\leqq r\leqq m}(-1)^r\binom{m}{r}x'^{m-r}\int_{x'-Y}^{x'+Y}u^{r-1}\,du \\ &= x'^m\log\left|\frac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right|+\sum_{1\leqq r\leqq m}(-1)^r\binom{m}{r}x'^{m-r}\frac{(x'+Y)^r-(x'-Y)^r}{r}. \end{aligned}$$

Thus in the strip $-1\leqq y'\leqq 1$

$$\begin{aligned} 2\pi D(x',y') &= \pi E(x',y')-i\sum_{m,n\geqq 0}a_{mn}x'^m y'^n\log\left|\frac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right| \\ &\quad -i\sum_{\substack{1\leqq r\leqq m\\ n\geqq 0}}a_{mn}x'^{m-r}y'^n(-1)^r\binom{m}{r}\frac{(x'+Y)^r-(x'-Y)^r}{r} \\ &= 2\pi D_0(x',y')\,E(x',y') \\ &\quad -i\sum_{\substack{1\leqq r\leqq m\\ n\geqq 0}}a_{mn}x'^{m-r}y'^n(-1)^r\binom{m}{r}\frac{(x'+Y)^r-(x'-Y)^r}{r}, \end{aligned} \tag{4·12}$$

where $D_0(x',y')$ is the value (2·35) of $D(x',y')$ for a true mirror. Outside the strip, $D(x',y')=0$ by (2·32).

(4·12), together with (2·24), gives the intensities seen under test on the disk and in the halo of a circular mirror whose error coefficients a_{mn} are defined by the equation (4·11).

As a special case suppose that in the expansion

$$\phi(x,y) = \sum_{m,n\geqq 0}\alpha_{mn}x^m y^n$$

of the error function $\phi(x,y)$ the squares of the quantities $\frac{2\pi}{\lambda}\alpha_{mn}$ are small enough to be neglected. Then on the mirror

$$E(x,y) = \exp\left[\frac{2\pi i}{\lambda}\phi(x,y)\right] = 1+\frac{2\pi i}{\lambda}\phi(x,y), \tag{4·13}$$

to a sufficient approximation, and the a_{mn} satisfy the equations*

$$a_{00} = 1+\frac{2\pi i}{\lambda}\alpha_{00}, \quad a_{mn} = \frac{2\pi i}{\lambda}\alpha_{mn} \quad (m^2+n^2>0). \tag{4·14}$$

The equations (4·12) and (2·24) then give

$$\begin{aligned} I(x',y') &= |\, 2\pi D(x',y')\,|^2 \\ &= \left|\left(\genfrac{}{}{0pt}{}{\pi}{0} - i\log\left|\frac{x'+Y}{x'-Y}\right|\right)\left(1+\frac{2\pi i}{\lambda}\phi(x',y')\right)\right. \\ &\qquad \left. +\frac{2\pi}{\lambda}\sum_{\substack{1\leqq r\leqq m\\ n\geqq 0}} \alpha_{mn}x'^{m-r}y'^{n}(-1)^r\binom{m}{r}\frac{(x'+Y)^r-(x'-Y)^r}{r}\right|^2 \\ &= \left(\genfrac{}{}{0pt}{}{\pi}{0}+\frac{2\pi}{\lambda}\phi(x',y')\log\left|\frac{x'+Y}{x'-Y}\right|+\frac{2\pi}{\lambda}\Sigma\right)^2+\left(\log\left|\frac{x'+Y}{x'-Y}\right|-\genfrac{}{}{0pt}{}{\pi}{0}\frac{2\pi}{\lambda}\Sigma\right)^2 \\ &= \genfrac{}{}{0pt}{}{\pi^2}{0}+\log^2\left|\frac{x'+Y}{x'-Y}\right|+\genfrac{}{}{0pt}{}{2\pi}{0}\frac{2\pi}{\lambda}\Sigma, \end{aligned}$$

on discarding terms in $(2\pi\alpha_{mn}/\lambda)^2$,

$$= I_0(x',y')+\genfrac{}{}{0pt}{}{4\pi^2/\lambda}{0}\sum_{\substack{1\leqq m\leqq r\\ n\geqq 0}}\alpha_{mn}x'^{m-r}y'^{n}(-1)^r\binom{m}{r}\frac{(x'+Y)^r-(x'-Y)^r}{r}, \tag{4·15}$$

where I_0 is the value (2·36) of the intensity for a true mirror and where the upper or lower alternatives are to be taken according as $x'^2+y'^2 \lessgtr 1$. It follows that *small errors on the mirror do not appreciably affect the halo*, a result which it is easy to derive directly from (2·32) and (4·13).

The further specialization

$$a_{00} = 1; \quad a_{20} = a_{02} = i(\alpha+\beta);$$

$$a_{40} = \tfrac{1}{2}a_{22} = a_{04} = -i\alpha; \quad \text{all other } a_{mn} = 0, \tag{4·16}$$

where α and β are small enough for their squares to be neglected, corresponds to the wave-function

$$E(x,y) = 1+i\alpha(r^2-r^4)+i\beta r^2. \tag{4·17}$$

---

* It makes no physical difference if it be supposed that $\alpha_{00} = 0$.

(4·12) and (4·15) now reduce to the equations*

$$2\pi D(x',y') = 2\pi D_0(x',y')\,E(x',y') - \tfrac{2}{3}\alpha x'\sqrt{(1-y'^2)}\,(2-3x'^2-5y'^2) - 2\beta x'\sqrt{(1-y'^2)}, \tag{4·18}$$

$$I(x',y') = I_0(x',y') - \genfrac{}{}{0pt}{}{2\pi}{0}\,[\tfrac{2}{3}\alpha x'\sqrt{(1-y'^2)}\,(2-3x'^2-5y'^2) + 2\beta x'\sqrt{(1-y'^2)}], \tag{4·19}$$

which therefore give the intensity seen on a mirror possessing a small amount of primary spherical aberration, tested with the knife-edge centrally set near mean focus.

4·2. In deriving formulae which shall predict the effects of varying the knife-edge setting, we treat lateral displacements of the knife-edge in a different way from variations in its focal setting. The difference in treatment corresponds to the circumstance that an optician using the Foucault test does in practice use these two adjustments of the knife-edge in different ways. Having chosen a focal setting which shows up as prominently as possible the error on which he is working, he tests by moving the knife-edge laterally to and fro and noting the changes in the appearance of the shadows.

If the errors on the mirror are small, the different focal settings which can be usefully employed are all covered, in the manner explained in § 2·3, by multiplying $E(x,y)$ by a factor $e^{i\alpha(x^2+y^2)}$ in which $\alpha = \frac{\pi}{\lambda s^2}\,\delta s$ is small. In the formula (4·13) $\phi(x,y)$ is replaced by $\phi(x,y)+\frac{\delta s}{2s^2}(x^2+y^2)$ and in (4·15) $\delta s/2s^2$ is added to α_{20} and α_{02}.

The situation is different with regard to the factor $e^{ic'x}$ which expresses the effect of a lateral shift of the knife-edge, since c' cannot be supposed small even in testing a true mirror, while values of c' up to 5 or 6 may be used in the observation of a fairly narrow zone only $\frac{1}{10}$ fringe in depth.

We therefore take the equation

$$2\pi D(x',y') = \genfrac{}{}{0pt}{}{\pi}{0}\,E(x',y') + i\int_{-\sqrt{(1-y'^2)}}^{\sqrt{(1-y'^2)}} e^{ic'(t-x')}\frac{E(t,y')}{t-x'}\,dt, \tag{4·21}$$

obtained on writing $e^{ic'x}E(x,y)$ for $E(x,y)$ and $e^{ic'x'}D(x',y')$ for $D(x',y')$ in (2·32), set

$$\begin{aligned} E(x,y) &= \sum_{m,n\geq 0} a_{mn}x^m y^n \quad (x^2+y^2\leqq 1) \\ &= 0 \quad (x^2+y^2>1), \end{aligned} \tag{4·22}$$

as before, and (supposing the errors of the mirror small compared with $\lambda/2\pi$) write

$$E(x,y) = 1+\frac{2\pi i}{\lambda}\left[\phi(x,y)+\frac{\delta s}{2s^2}(x^2+y^2)\right] \tag{4·23}$$

---

* In an earlier version of the present paper, these equations were obtained from an extension of the analysis of Zernike. I was not able, however, to obtain (4·26) in this way.

to express the effects of an increase δs in the focal distance s of the knife-edge from the mirror. (4·21) then becomes

$$2\pi D(x',y') = \genfrac{}{}{0pt}{}{\pi}{0} E(x',y') - i\int_{x'-\surd(1-y'^2)}^{x'+\surd(1-y'^2)} E(x'-u,y')\,e^{ic'u}\frac{du}{u}$$

$$= \genfrac{}{}{0pt}{}{\pi}{0} E(x',y') - i \sum_{m,n\geqq 0} a_{mn}y'^n \sum_{0\leqq r\leqq m} (-1)^r \binom{m}{r} x'^{m-r} I_{r-1},$$

where $$I_{r-1} = \int_{x'-Y}^{x'+Y} u^{r-1}e^{-ic'u}\,du \quad (r\geqq 0)$$

and Y is written for $\sqrt{(1-y'^2)}$.

Here

$$I_{-1} = \mathrm{Ci}\{|\,c'(x'+Y)\,|\} + \mathrm{Ci}\{|\,c'(x'-Y)\,|\} - i\mathrm{Si}\{c'(x'+Y)\} + i\mathrm{Si}\{c'(x'-Y)\},$$

as in § 3·1, while for $r\geqq 1$

$$I_{r-1} = \sum_{S=1}^{r} (-1)^{S-1}\left(\frac{i}{c'}\right)^S \frac{(r-1)!}{(r-S)!}\left[(x'+Y)^{r-S}e^{-ic'(x'+Y)} - (x'-Y)^{r-S}e^{-ic'(x'-Y)}\right].$$

When $c'=0$, these expressions are to be interpreted as their limiting values; for example, I_{-1} becomes $\log\left|\dfrac{x'+\sqrt{(1-y'^2)}}{x'-\sqrt{(1-y'^2)}}\right|$. For general c', it follows that

$$\begin{aligned}2\pi D(x',y') &= \genfrac{}{}{0pt}{}{\pi}{0} E(x',y') - i\sum_{m,n\geqq 0} a_{mn}x'^m y'^n I_{-1}\\ &\qquad - ie^{-ic'x'} \sum_{\substack{n\geqq 0\\ 1\leqq S\leqq r\leqq m}} a_{mn}x'^{m-r}y'^n(-1)^{r+S-1}\binom{m}{r}\left(\frac{i}{c'}\right)^S \frac{(r-1)!}{(r-S)!} M(r-S)\\ &= 2\pi D_0(x',y')\,E(x',y')\\ &\qquad - ie^{-ic'x'} \sum_{\substack{n\geqq 0\\ 1\leqq S\leqq r\leqq m}} a_{mn}x'^{m-r}y'^n(-1)^{r-S+1}\binom{m}{r}\left(\frac{i}{c'}\right)^S \frac{(r-1)!}{(r-S)!} M(r-S),\end{aligned} \tag{4·24}$$

where

$$\begin{aligned}2\pi D_0(x',y') &= \genfrac{}{}{0pt}{}{\pi}{0} - iI_{-1}\\ &= \genfrac{}{}{0pt}{}{\pi}{0} - i\mathrm{Ci}\{|\,c'[x'+\sqrt{(1-y'^2)}]\,|\} + i\mathrm{Ci}\{|\,c'[x'-\sqrt{(1-y'^2)}]\,|\}\\ &\qquad - \mathrm{Si}\{|\,c'[x'+\sqrt{(1-y'^2)}]\,|\} + \mathrm{Si}\{c'[x'-\sqrt{(1-y'^2)}]\}\end{aligned}$$

is the value (3·2) of $2\pi D(x',y')$ for a true mirror (that is, one with $E(x,y)=1$ or 0) at the same lateral setting of the knife-edge and

$$M(r-S) = [x'+\sqrt{(1-y'^2)}]^{r-S}\,e^{-ic'\surd(1-y'^2)} - [x'-\sqrt{(1-y'^2)}]\,e^{ic'\surd(1-y'^2)}.$$

Writing for shortness

$$\frac{\sin[c'\sqrt{(1-y'^2)}]}{c'\sqrt{(1-y'^2)}} = S, \quad \cos[c'\sqrt{(1-y'^2)}] = C,$$

we find

$$\begin{aligned} M(0) &= -2ic'YS, \\ M(1) &= x'M(0)+2YC, \\ M(2) &= (x'^2+Y^2)M(0)+4x'YC, \\ M(3) &= (x'^3+3x'Y^2)M(0)+2(3x'^2+Y^2)YC. \end{aligned} \tag{4·25}$$

Primary spherical aberration of small amount, tested near mean focus

Suppose that in (4·24) all the a_{mn} are zero except $a_{00}, a_{20}, a_{02}, a_{40}, a_{22}, a_{04}$. Then

$$2\pi D(x',y') = 2\pi D_0(x',y')\,E(x',y') - ie^{-ic'x'}\Sigma,$$

where

$$\begin{aligned} \Sigma &\doteq \sum_{\substack{1\leqq S\leqq r\leqq m\\ n\geqq 0}} a_{mn}x'^{m-r}y'^{n}(-1)^{r-S+1}\left(\frac{i}{c'}\right)^{S}\frac{(r-1)!}{(r-S)!}M(r-S) \\ &= (a_{20}+a_{22}y'^2)\left[-M(0)\sum_{r=1}^{2}x'^{2-r}\left(\frac{i}{c'}\right)^{r}\binom{2}{r}+M(1)\frac{i}{c'}\right] \\ &\quad +a_{40}\Bigg[-M(0)\sum_{r=1}^{4}x'^{4-r}(r-1)!\left(\frac{i}{c'}\right)^{r}\binom{4}{r}+M(1)\sum_{r=2}^{4}x'^{4-r}(r-1)!\left(\frac{i}{c'}\right)^{r-1}\binom{4}{r} \\ &\qquad\qquad -M(2)\sum_{r=3}^{4}x'^{4-r}\left(\frac{i}{c'}\right)^{r-2}\frac{(r-1)!}{2!}\binom{4}{r}+M(3)\frac{i}{c'}\Bigg] \\ &= (a_{20}+a_{22}y'^2)\left[-2x'YS+\frac{2iY}{c'}(C-S)\right] \\ &\quad +a_{40}\left[-2x'YS(x'^2+Y^2)+\frac{2iY}{c'}(C-S)\left(x'^2+\frac{2x'}{c'}i-\frac{6}{c'^2}+Y^2\right)-\frac{2iY}{c'}\,.\,2SY^2\right]. \end{aligned} \tag{4·26}$$

The intensity distribution for a mirror possessing a small amount of primary spherical aberration, tested with the knife-edge near mean focus but not necessarily central, is now obtained on giving the six coefficients $a_{00}, a_{20}, a_{02}, a_{40}, a_{22}, a_{04}$ the values (4·16).

5. THEORETICAL SENSITIVITY OF THE KNIFE-EDGE TEST

The results of the last section can be used to estimate the theoretical sensitiveness of the Foucault test for errors of different types. Only the two simplest cases are considered here, namely, the determination of the focus of an aberration-free pencil and the detection of primary spherical aberration, and it is supposed that the knife-edge intersects the axis of the system. In practice, the lateral adjustment of the knife-edge would of course be varied.

In the first case, the effect of an error δs of excess in the determination of the focal distance is an apparent error $r^2\delta s/16F^2$ on the surface of the mirror on which, as in § 2, the errors may be supposed to be located. F denotes the aperture ratio of the pencil, or of this mirror, and $d(r) = r^2\delta s/8F^2$ measures the variation in optical path distance from points on a wave-front just leaving this mirror to the axial point of the knife-edge. At the edge ($r = 1$) of the pencil, the variation reaches its greatest value $\delta s/8F^2$. Thus the maximum variation will be $\frac{1}{20}\lambda$ if $\delta s = \frac{2}{5}\lambda F^2$, $d(r) = \frac{1}{20}\lambda r^2$. For an $f/5$ pencil this corresponds to a focusing error $\delta s = 0{\cdot}0002$ in.

For such small path differences $d(r)$, the approximation

$$\exp\left[\frac{2\pi i}{\lambda}d(r)\right] = 1+\frac{2\pi i}{\lambda}d(r) = 1+\frac{\pi i}{10}r^2 \tag{5·1}$$

can be used without introducing more than a few per cent of variation in the amplitude of the wave leaving the mirror. The expression for the intensity I_1 is then obtained by taking $\alpha = 0$, $\beta = \frac{1}{10}\pi$ in (4·19). On the disk its value is

$$I_1 = I_0-\tfrac{2}{5}\pi^2x'\sqrt{(1-y'^2)}. \tag{5·2}$$

Figure 8 shows the isophotal lines in this case, and figure 9 the intensity distribution along the horizontal diameter as calulated from (5·2) and also as calculated from the more exact formulae (4·12), (4·13) and (2·24).*

In the second case, namely, primary spherical aberration, the retardation function at mean focus has the form $d(r) = Ar^2(1-r^2)$. To obtain a maximum variation of optical path distance of $\frac{1}{20}\lambda$, we take $A = \frac{1}{5}\lambda$ and obtain, to a sufficient approximation,

$$\exp\left[\frac{2\pi i}{\lambda}d(r)\right] = 1+\frac{2\pi i}{\lambda}d(r) = 1+\frac{2\pi i}{5}r^2(1-r^2). \tag{5·3}$$

The intensity I_2 is then obtained from (4·19) by taking $\alpha = \frac{2}{5}\pi$, $\beta = 0$, and its value on the disk is

$$I_2 = I_0-\frac{8\pi^2}{15}x'\sqrt{(1-y'^2)}\,(2-3x'^2-5y'^2). \tag{5·4}$$

Figure 10 shows the isophotal lines in this case and figure 11 the intensity distribution along the horizontal diameter as calculated from (5·4) and from the more exact formulae (4·12), (4·13) and (2·24).†

To draw numerical conclusions about the limit of sensitivity of the knife-edge test in these two cases, some assumptions must be made about the least departure from symmetry in the intensity distribution which can be detected by visual inspection. If it be assumed that a 40 % difference in intensity between two fairly large, symmetrically disposed areas of the disk is easily detectable, then one can say that an error of focus which results in optical path differences of $\frac{1}{20}$ fringe is easily detectable by the knife-edge test, and that the same is true of $\frac{1}{20}$ fringe of primary spherical aberration tested at mean focus. If it be assumed that the observer can

---

* The error of approximation in (4·13) was reduced by using for the more exact calculation the retardation function $d(r) = \frac{1}{20}\lambda(r^2-\frac{1}{2})$.

† The retardation-function was taken as $\frac{\lambda}{5}(r^2-r^4-\frac{1}{6})$ for the more exact calculation.

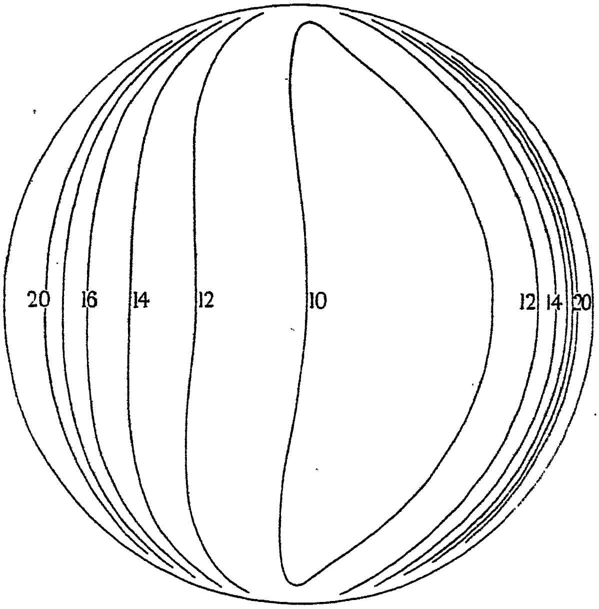


FIGURE 8. Disk-intensities under knife-edge test of a true mirror with $\frac{1}{20}$ fringe focus-error; $d(r) = \frac{1}{20}\lambda r^2$.

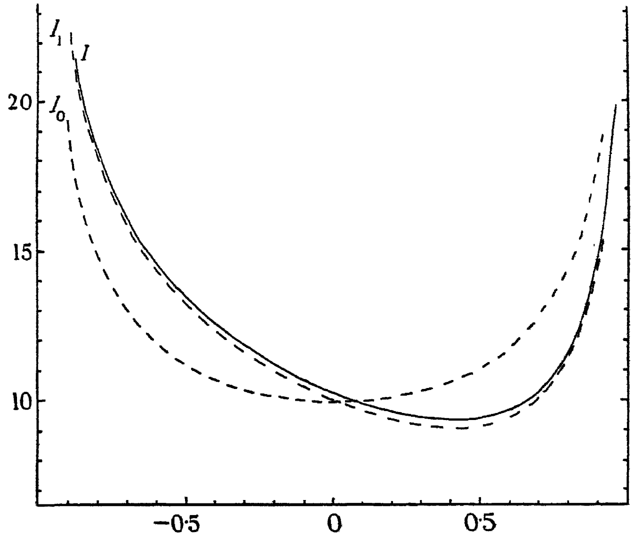


FIGURE 9. Intensities along horizontal diameter of a mirror with $\frac{1}{20}$ fringe of focus error, tested with knife-edge centrally set. I = intensity calculated from (4·12), (4·13); I_1 = intensity calculated from the approximate formula (5·2); I_0 = intensity for a true mirror.

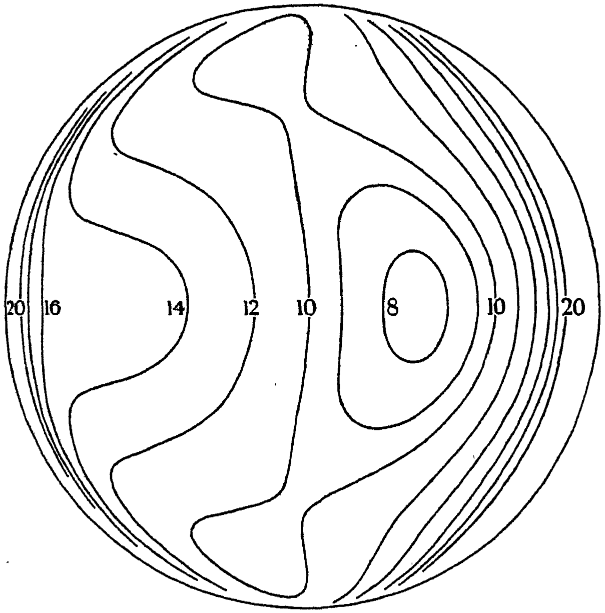


FIGURE 10. Disk-intensities under knife-edge test of a mirror with $\frac{1}{20}$ fringe primary spherical aberration: $d(r) = \frac{1}{5}\lambda r^2(1-r^2)$.

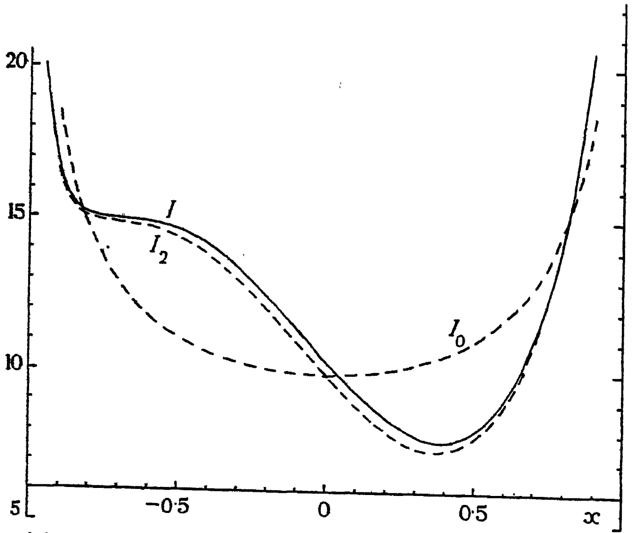


FIGURE 11. Intensities along horizontal diameter of a mirror with $\frac{1}{10}$ fringe of primary spherical aberration, tested with knife-edge at mean focus. I = intensity calculated from (4·12), (4·13); I_2 = intensity calculated from the approximate formula (5·4); I_0 = intensity for a true mirror.

detect with certainty a 20 % difference of intensity, the corresponding conclusion is that he can detect with certainty $\frac{1}{40}$ fringe of error of either of these two types. The calculation is simplified by observing that, for small errors, $I-I_0$ is linear in the coefficients α, β; consequently its values for amounts up to $\frac{1}{20}$ fringe of spherical aberration or of focus error, or of a mixture of the two, can be deduced by linear interpolation from the cases already discussed (figures 7*c*, 8 and 10).

As a rule, a sufficiently good idea of the knife-edge shadows can be got from the intensity distribution along the horizontal diameter. In table 1 the intensity changes along this diameter caused by amounts $\frac{1}{20}\lambda$ of focus error and of primary spherical aberration are given, together with the intensity values for a true mirror. The table enables the intensity distribution to be calculated for small amounts of primary spherical aberration, tested at different axial settings of the knife-edge near best focus.

Figures 12 and 13 show the intensities calculated in this way for mirrors with $\frac{1}{20}$ fringe and $\frac{1}{40}$ fringe of primary spherical aberration, tested with the knife-edge on axis at equidistant focal settings ranging from paraxial focus (curve *A*) through mean focus (curve *E*). Curves *A* and *B* of figure 12, in which extrapolation is pushed farthest, are rather rough approximations to the true values of the intensities. It is of interest that the illumination near the centre of the disk appears most uniform, not when the knife-edge is at paraxial focus, as ray theory would predict, but when it is about midway between paraxial and mean focus.

TABLE 1

I_0 = intensity for a true mirror; I_1-I_0 = intensity change for $\frac{1}{20}\lambda$ of focus error; I_2-I_0 = intensity change for $\frac{1}{20}\lambda$ of primary spherical aberration.

x	I_0	I_1-I_0	I_2-I_0
0	9·870	0	0
0·1	9·910	−0·395	−1·037
0·2	10·034	−0·790	−1·979
0·3	10·253	−1·184	−2·732
0·4	10·587	−1·579	−3·200
0·5	11·077	−1·974	−3·290
0·6	11·791	−2·369	−2·906
0·7	12·878	−2·763	−1·953
0·8	14·697	−3·158	−0·337
0·9	18·539	−3·553	+2·037

In the estimate made above of the sensitivity of the Foucault test as a means of detecting primary spherical aberration, it was taken for granted that the knife-edge could be accurately located at mean focus. In practice this is not the case; the 'focal setting' of the knife-edge is tacitly defined as that which makes the disk intensities satisfy some condition of left-to-right symmetry. When the mirror errors are of the order of a fringe or more, or when fairly rapid zonal error is present, a convenient condition is that the right and left halves of the disk shall become completely dark simultaneously. This is analogous to defining the position of best focus by means of

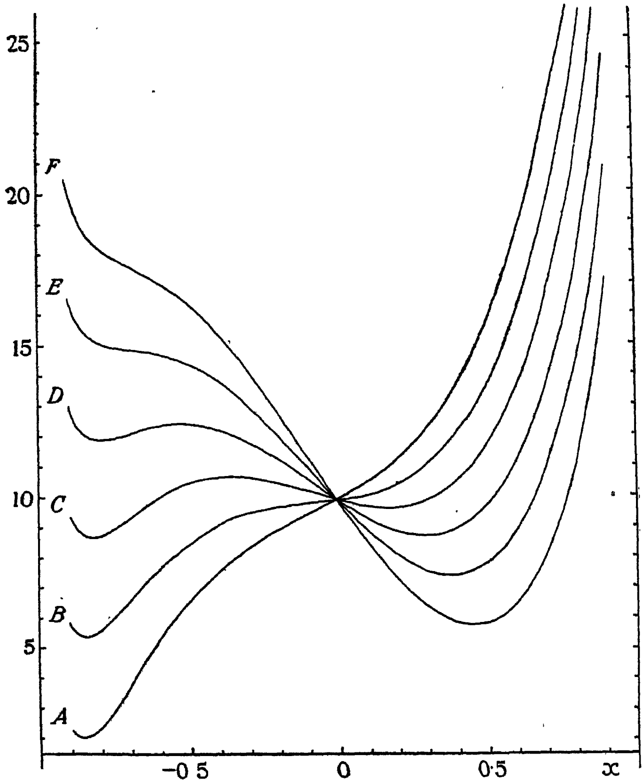


FIGURE 12

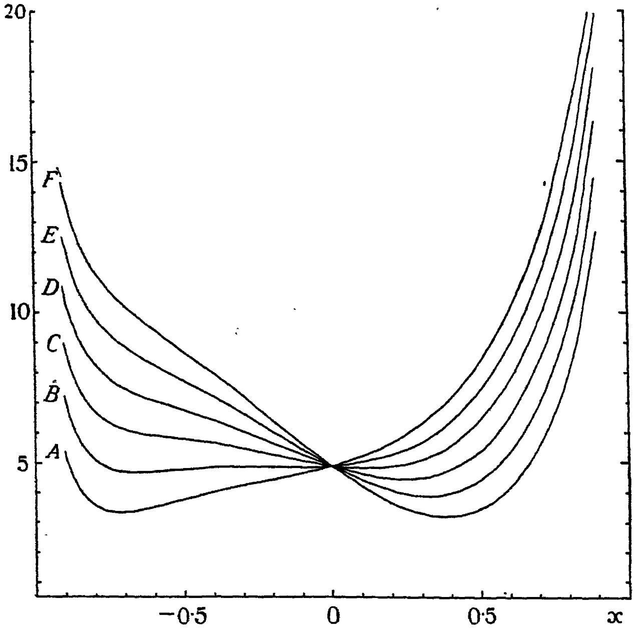


FIGURE 13

the circle of least confusion. When, however, the errors are slow and very small, the dominance of the bright rim defeats this method, and it is more practicable to use a different condition, namely, that when the knife-edge bisects the image the total illumination in the two halves of the disk shall be judged equal.

To estimate the sensitivity of the test in a way which shall have a practical meaning, one must therefore consider a whole range of settings near mean focus. This can be done, in the two cases already discussed, by means of figures 12 and 13, provided we are content to consider only intensities along the main diameter. From the practical point of view this is not a very serious restriction. Figures 12 and 13 show that the selected setting will in both cases lie between those corresponding to curve C and to curve E, i.e. a little inside mean focus. Applying the above sensitivity criteria to the intensities along the horizontal diameter (which are taken to represent those over the whole disk sufficiently well for this purpose), we conclude that an error of $\frac{1}{20}$ fringe of primary spherical aberration can still be detected with ease, and one of $\frac{1}{40}$ fringe with certainty, in the case where the focal setting has to be chosen empirically.

REFERENCES

Banerji, S. 1918 *Astrophys. J.* **48**, 50–58.
Banerji, S. 1919 *Phil. Mag.* **37**, 114 and plate 3, figure 9.
Couder, A. 1937 *Rev. Opt.* **16**, 17.
Ellison, W. F. A. 1939 *Amateur telescope making*, pp. 83 and 96.
Foucault, L. 1859 *Ann. Obs. Imp. Paris*, **5**, 197–237.
Gascoigne, S. C. G. 1945 *Mon. Not. R. Ast. Soc.* **104**, 326–334.
Michelson, A. A. 1905 *Phil. Mag.* (6), **9**, 506.
Poincaré, H. 1892 *Acta Math.* **16**, 297–339.
Poincaré, H. 1897 *Acta Math.* **20**, 313–356.
Porter, R. W. 1939 *Amateur telescope making*, p. 8.
Rayleigh 1917 *Phil. Mag.* (6), **33**, 161–178.
Zernike, F. 1934 *Physica*, **1**, 689–704.

An integral-equation method of solving plane diffraction problems

By E. T. Copson

(*Communicated by Sir Edmund Whittaker, F.R.S.—Received* 16 *April* 1945)

1. Introduction

When monochromatic electromagnetic waves with electric and magnetic vectors* $\mathbf{E}^i$, $\mathbf{H}^i$ are incident on a perfectly reflecting screen S, the waves are scattered at the surface of the screen. The problem of reflexion and diffraction consists in determining the scattered field $\mathbf{E}^s$, $\mathbf{H}^s$, which satisfies the following conditions:

(i) it is a solution of Maxwell's equations

$$ik\mathbf{E} = \operatorname{curl}\mathbf{H}, \quad ik\mathbf{H} = -\operatorname{curl}\mathbf{E}, \quad \operatorname{div}\mathbf{E} = 0, \quad \operatorname{div}\mathbf{H} = 0;$$

(ii) on S, $$(\mathbf{E}^i+\mathbf{E}^s)\times\mathbf{N} = 0, \quad (\mathbf{H}^i+\mathbf{H}^s).\mathbf{N} = 0,$$

where $\mathbf{N}$ is the outward normal unit vector to S;

(iii) it satisfies Sommerfeld's radiation condition at infinity.

There is only one field $\mathbf{E}^s$, $\mathbf{H}^s$ which satisfies these conditions.

The simplest form of the radiation condition is to assume that $k = p-iq$, where p and q are positive and q is small, and to require each field component to vanish at infinity. This corresponds physically to assuming that the medium has a small conductivity. The imaginary part of k can be made zero at the end of the analysis. It is important to note that a field which satisfies the radiation condition and which has no singularity anywhere in space is null.

If V is a rectangular cartesian component of $\mathbf{E}^s$ or of $\mathbf{H}^s$, it satisfies the wave-equation

$$\nabla^2 V + k^2 V = 0$$

and the radiation condition, and is therefore completely determined by the values on S either of V or of its normal derivative. The explicit formulae are

$$V(P) = \frac{1}{4\pi}\iint_S V(P')\frac{\partial G_1(P,P')}{\partial N'}\,dS, \tag{1·1}$$

$$V(P) = -\frac{1}{4\pi}\iint_S \frac{\partial V(P')}{\partial N'}\,G_2(P,P')\,dS, \tag{1·2}$$

where $\mathbf{N}'$ is the outward normal unit vector to S at the integration point P', and G_1 and G_2 are the Green's functions of the first and second kinds. These formulae have been little used in diffraction theory† for two reasons. First, if V is a component of $\mathbf{E}^s$ or $\mathbf{H}^s$, the boundary conditions do not determine V or $\partial V/\partial N$ on S. Secondly, the determination of the Green's functions is difficult and is equivalent to the solution of a scalar diffraction problem.

---

* Here and throughout the paper, a time factor e^{iwt} $(w>0)$ is understood.

† The only case I know is Schwarzschild's work (1902) on diffraction by a slit.

It was for the first of these reasons that recourse has been had to vector solutions of Maxwell's equations, such as the Larmor-Tedone formulae (Baker & Copson 1939). But these formulae are no better suited to the solution of diffraction problems, since they require a complete knowledge of the scattered field at the surface of the screen. There is no special merit in a vector solution; for if the correct boundary values could be inserted in (1·1) or (1·2), the resulting expressions for $\mathbf{E}^s$ and $\mathbf{H}^s$ would satisfy Maxwell's equations automatically, since these boundary value problems have unique solutions. Unfortunately it is not, in general, possible to do this.

The problem of reflexion by a perfectly conducting plane is quite exceptional. If the plane is $z = 0$ and if P lies in $z > 0$, (1·1) and (1·2) simplify (Rayleigh 1896) to

$$V(x,y,z) = -\frac{1}{2\pi}\frac{\partial}{\partial z}\iint V(x',y',z')\,\phi\,dx'\,dy', \tag{1·3}$$

$$V(x,y,z) = -\frac{1}{2\pi}\iint \frac{\partial V(x',y',z')}{\partial z'}\,\phi\,dx'\,dy', \tag{1·4}$$

where integration is over the plane $z' = 0$ and $\phi = e^{-ikR}/R$, with

$$R^2 = (x-x')^2+(y-y')^2+(z-z')^2.$$

When $z = 0$, we have $\quad E^s_x = \mathscr{E}_x,\quad E^s_y = \mathscr{E}_y,\quad H^s_z = \mathscr{H}_z,$

where $\mathscr{E}_x$, $\mathscr{E}_y$, $\mathscr{H}_z$ are functions of (x,y) known from the boundary conditions, and so (1·3) gives the components E^s_x, E^s_y, H^s_z of the reflected field in the half-space $z > 0$. Again, by Maxwell's equations, one can express the values on $z = 0$ of $\partial H^s_x/\partial z$, $\partial H^s_y/\partial z$, $\partial E^s_z/\partial z$ in terms of $\mathscr{E}_x$ and $\mathscr{E}_y$, and (1·4) will give the remaining field components in $z > 0$. A complete solution of the problem of reflexion of electromagnetic waves by a perfectly conducting plane $z = 0$ is thus provided by the scalar formulae (1·3) and (1·4). A similar procedure will solve the reflexion problem in $z < 0$, provided that the signs of the expressions on the right-hand sides of (1·3) and (1·4) are changed.

Now consider the problem when the screen lies in the plane $z = 0$ but does not cover the whole plane. It has one or more holes S_1, say; the metal of the screen will be denoted by S_2, so that S_1 and S_2 fill up the whole plane. The boundary values of E^s_x, E^s_y, H^s_z, $\partial H^s_x/\partial z$, $\partial H^s_y/\partial z$, $\partial E^s_z/\partial z$ are the same as before on S_2, but are unknown on S_1. Equations (1·3) and (1·4) will still give the scattered field in $z > 0$ but will involve certain functions unknown on S_1; and similarly in $z < 0$. The continuity of the field through the holes S_1 gives a set of integral equations to determine these unknown functions. These integral equations are found below, and turn out to be of an unusual type. They are, in effect, of the form

$$\iint f(x',y')\,K(x,y;x',y')\,dx'\,dy' = F(x,y),$$

where integration is over the whole plane; the nucleus K is singular, $f(x,y)$ is zero on S_2 and unknown on S_1, and $F(x,y)$ is unknown on S_2 but known on S_1.

This formulation of the problem of diffraction by a perfectly conducting plane screen is equivalent to assuming that the scattered field is due to a magnetic charge and current sheet over the holes S_1 in the screen, a device which has recently been used successfully by Bethe (1944) in connexion with diffraction by a small circular hole.

By using properties of the scattered field revealed in this new formulation of the diffraction problem, a second set of integral equations involving functions unknown on the screen S_2 is obtained. This second formulation of the problem is equivalent to assuming that the scattered field is due to electric charges and currents induced in the screen by the incident field, an idea originally suggested by Poincaré and recently used by W. Magnus (1941) is his solution of Sommerfeld's diffraction problem.

Section 5 shows how the two methods may be applied to the Sommerfeld problem of the diffraction of plane polarized waves by a perfectly conducting half-plane. They can also be used to solve approximately the problems of diffraction by a narrow slit, a narrow blade, a small circular disk or a small circular hole. All these problems were solved by Rayleigh (1897) nearly fifty years ago for incident plane waves; the last has recently been solved by Bethe* for a general incident field. Both authors make considerable use of physical arguments. Bethe uses a fictitious magnetic charge and current sheet. Rayleigh, on the other hand, in dealing with diffraction by a small ellipsoid (of which a circular disk is a limiting form) assumes that the field near the obstacle can be calculated by the methods of electrostatics and magnetostatics; the field at large distances is then deduced by comparing the field near the obstacle with that due to a Hertzian oscillator. The methods developed here make no appeal to physical intuition.

The two formulations of the problem of diffraction by a plane screen are complementary, in the sense that one involves functions unknown on the screen, the other functions unknown on the gaps in the screen. It is shown in § 7 that this leads at once to the rigorous form of Babinet's principle recently enunciated by H. G. Booker.

2. THE FUNDAMENTAL FORMULAE

It was seen in § 1 that, under certain conditions, an electromagnetic field can be expressed in terms of the values on the plane $z = 0$ of the two tangential components of **E** and the normal component of **H**. Theorem 1 shows this dependence on the boundary values explicitly.

THEOREM 1. *Let* **E**, **H** *be an electromagnetic field which is regular in* $z \geqslant 0$ *and satisfies the radiation condition in* $z > 0$. *Let* $\mathscr{E}_x$, $\mathscr{E}_y$, $\mathscr{H}_z$ *denote the values of* E_x, E_y, H_z *at the boundary point* $(x', y', 0)$ *and let*

$$\mathscr{E}_x e^{-ikr'}, \quad \mathscr{E}_y e^{-ikr'}, \quad \mathscr{H}_z e^{-ikr'}$$

---

* Bethe (1944) seems to be unaware of Rayleigh's work; for he remarks that 'the available theoretical methods are entirely inadequate for the treatment of our problem'.

tend to zero as $r' = \sqrt{(x'^2+y'^2)} \to \infty$, *uniformly in all directions. Then the field in the half-space* $z > 0$ *is given by*

$$\left.\begin{aligned} E_x &= -\frac{\partial U}{\partial z}, & E_y &= -\frac{\partial V}{\partial z}, & E_z &= \frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}, \\ H_x &= -ikV-\frac{\partial W}{\partial x}, & H_y &= ikU-\frac{\partial W}{\partial y}, & H_z &= -\frac{\partial W}{\partial z}, \end{aligned}\right\} \tag{2·1}$$

where

$$U = \frac{1}{2\pi}\iint \mathscr{E}_x\phi\, dx'\,dy', \quad V = \frac{1}{2\pi}\iint \mathscr{E}_y\phi\, dx'\,dy', \quad W = \frac{1}{2\pi}\iint \mathscr{H}_z\phi\, dx'\,dy'. \tag{2·2}$$

The integrals are extended over the whole (x', y') *plane, and*

$$\phi = \frac{1}{R}e^{-ikR},$$

where $$R^2 = (x-x')^2+(y-y')^2+z^2.$$

The functions U, V, W *are connected by the relation*

$$ikW = -\frac{\partial V}{\partial x}+\frac{\partial U}{\partial y}. \tag{2·3}$$

By saying that the field is regular in $z \geqslant 0$, it is meant that all the components of **E** and **H** and their first and second partial derivatives are continuous in $z \geqslant 0$. This implies that $\mathscr{E}_x$, $\mathscr{E}_y$, $\mathscr{H}_z$ are continuous functions of (x', y'), a restriction which will have to be removed later. The radiation condition is assumed to hold only in $z > 0$, and lighter conditions are applied to the boundary values.

Under the conditions of the theorem, (1·3) and (1·4) can be applied to each rectangular cartesian component of **E** and **H**, and the formula for E_x, E_y, H_z follow at once. Next, by Maxwell's equations, it follows that

$$\frac{\partial H_x}{\partial z} = ikE_y+\frac{\partial H_z}{\partial x},$$

so that

$$\begin{aligned} H_x &= -\frac{1}{2\pi}\iint\left\{ik\mathscr{E}_y+\frac{\partial \mathscr{H}_z}{\partial x'}\right\}\phi\, dx'\,dy' \\ &= -ikV-\frac{1}{2\pi}\iint\left\{\frac{\partial}{\partial x'}(\mathscr{H}_z\phi)-\mathscr{H}_z\frac{\partial\phi}{\partial x'}\right\}dx'\,dy'. \end{aligned}$$

But the integral of $\partial(\mathscr{H}_z\phi)/\partial x'$ over the whole plane vanishes by Green's theorem, since $\mathscr{H}_z e^{-ikr'}$ tends to zero as $r' \to \infty$ uniformly in all directions. Hence

$$\begin{aligned} H_x &= -ikV+\frac{1}{2\pi}\iint \mathscr{H}_z\frac{\partial\phi}{\partial x'}\,dx'\,dy' = -ikV-\frac{1}{2\pi}\iint \mathscr{H}_z\frac{\partial\phi}{\partial x}\,dx'\,dy' \\ &= -ikV-\frac{\partial W}{\partial x}. \end{aligned}$$

The formulae for H_y, E_z can be obtained in a similar way. The formula (2·3) is also a consequence of Green's theorem and Maxwell's equations. It should be noted that if U, V, W are any solutions of the wave-equation, the expressions (2·1) satisfy Maxwell's equations if, and only if, U, V, W satisfy (2·3).

The corresponding theorem for the half-space $z < 0$ is

THEOREM 2. *Let* **E**, **H** *be an electromagnetic field which is regular in* $z \leqslant 0$ *and satisfies the radiation condition in* $z < 0$. *Let* $\mathscr{E}_x$, $\mathscr{E}_y$, $\mathscr{H}_z$ *denote the values of* E_x, E_y, H_z *at the boundary point* $(x', y', 0)$, *and let*

$$\mathscr{E}_x e^{-ikr'}, \quad \mathscr{E}_y e^{-ikr'}, \quad \mathscr{H}_z e^{-ikr'}$$

tend to zero as $r' \to \infty$, *uniformly in all directions. Then the field in* $z < 0$ *is given by*

$$\left.\begin{aligned} E_x &= \frac{\partial U}{\partial z}, & E_y &= \frac{\partial V}{\partial z}, & E_z &= -\frac{\partial U}{\partial x} - \frac{\partial V}{\partial y}, \\ H_x &= ikV + \frac{\partial W}{\partial x}, & H_y &= -ikU + \frac{\partial W}{\partial y}, & H_z &= \frac{\partial W}{\partial z}, \end{aligned}\right\} \tag{2·4}$$

where U, V, W *are defined by* (2·2) *and are connected by* (2·3).

One would expect that the field defined by (2·4) would be the analytical continuation of that defined by (2·1), but this is not so. For if it were, (2·1) and (2·4) would define a field regular in all space and satisfying the radiation condition; it would therefore be a null field which is certainly not the case. The explanation is, as will be seen in § 3, that (2·1) and (2·4) define a field due to a magnetic charge and current sheet in the plane $z = 0$.

It will be necessary later to use the fact that E_x, E_y, H_z *defined by* (2·1) *and* (2·4) *are even functions of* z, *whereas* H_x, H_y, E_z *are odd functions of* z.

Since Maxwell's equations are invariant under the transformation $\mathbf{E} \to \mathbf{H}$, $\mathbf{H} \to -\mathbf{E}$, a second set of fundamental formulae involving the boundary values of H_x, H_y, E_z can be written down without proof.

THEOREM 3. *Let* **E**, **H** *be an electromagnetic field which is regular in* $z \geqslant 0$ *and satisfies the radiation condition in* $z > 0$. *Let* $\mathscr{H}_x$, $\mathscr{H}_y$, $\mathscr{E}_z$ *denote the values of* H_x, H_y, E_z *at the boundary point* $(x', y', 0)$, *and let*

$$\mathscr{H}_x e^{-ikr'}, \quad \mathscr{H}_y e^{-ikr'}, \quad \mathscr{E}_z e^{-ikr'}$$

tend to zero as $r' \to \infty$, *uniformly in all directions. Then the field in the half-space* $z > 0$ *is given by*

$$\left.\begin{aligned} E_x &= ikV' - \frac{\partial W'}{\partial x}, & E_y &= -ikU' - \frac{\partial W'}{\partial y}, & E_z &= -\frac{\partial W'}{\partial z}, \\ H_x &= -\frac{\partial U'}{\partial z}, & H_y &= -\frac{\partial V'}{\partial z}, & H_z &= \frac{\partial U'}{\partial x} + \frac{\partial V'}{\partial y}, \end{aligned}\right\} \tag{2·5}$$

where

$$U' = \frac{1}{2\pi}\iint \mathscr{H}_x \phi \, dx' \, dy', \quad V' = \frac{1}{2\pi}\iint \mathscr{H}_y \phi \, dx' \, dy', \quad W' = \frac{1}{2\pi}\iint \mathscr{E}_z \phi \, dx' \, dy'. \tag{2·6}$$

The functions U', H', W' *are connected by the relation*

$$ikW' = \frac{\partial V'}{\partial x} - \frac{\partial U'}{\partial y}. \tag{2·7}$$

The corresponding formulae for the half-space $z<0$ are

$$\left.\begin{aligned} E_x &= -ikV'+\frac{\partial W'}{\partial x}, & E_y &= ikU'+\frac{\partial W'}{\partial y}, & E_z &= \frac{\partial W'}{\partial z}, \\ H_x &= \frac{\partial U'}{\partial z}, & H_y &= \frac{\partial V'}{\partial z}, & H_z &= -\frac{\partial U'}{\partial x}-\frac{\partial V'}{\partial y}. \end{aligned}\right\} \tag{2·8}$$

The field defined by (2·8) is not the analytical continuation of the field defined by (2·5). The physical reason is that both are due to an electric charge and current sheet in the plane $z=0$.

When these theorems are applied to diffraction problems, the boundary values will not be continuous. For example, in the applications of theorem 1, $\mathscr{E}_x$, $\mathscr{E}_y$, $\mathscr{H}_z$ will be discontinuous on certain curves which are the boundaries of the metal screen. The formulae of the theorem will, however, still hold provided that the integrals (2·2) converge; E_x, E_y, H_z will tend respectively to $\mathscr{E}_x$, $\mathscr{E}_y$, $\mathscr{H}_z$ as (x,y,z) tends to any boundary point $(x', y', +0)$ at which these functions are continuous.

3. CONNEXION WITH THE LARMOR-TEDONE FORMULAE

The formulae of § 2 may also be deduced from the Larmor-Tedone vector solution of Maxwell's equations, which has the advantage of showing the physical meaning of the formulae. The Larmor-Tedone vector solution (Baker & Copson 1939) of the boundary value problem for the half-space $z>0$ is

$$\mathbf{E} = -\operatorname{grad}\Phi - ik\mathbf{A} - \operatorname{curl}\mathbf{B}, \quad \mathbf{H} = -\operatorname{grad}\Psi - ik\mathbf{B} + \operatorname{curl}\mathbf{A}, \tag{3·1}$$

where Φ and $\mathbf{A}$ are the scalar and vector potentials of an electric charge and current sheet in the plane $z=0$, whereas Ψ and $\mathbf{B}$ are the scalar and vector anti-potentials of a magnetic charge and current sheet. When $z<0$, the expressions on the right-hand sides of equations (3·1) are identically zero. In the notation of § 2, the explicit formulae for the potentials are

$$\left.\begin{aligned} \Phi &= \frac{1}{4\pi}\iint \mathscr{E}_z\phi\,dx'\,dy', & \mathbf{A} &= \frac{1}{4\pi}\iint (\mathbf{j}\mathscr{H}_x-\mathbf{i}\mathscr{H}_y)\,\phi\,dx'\,dy', \\ \Psi &= \frac{1}{4\pi}\iint \mathscr{H}_z\phi\,dx'\,dy', & \mathbf{B} &= -\frac{1}{4\pi}\iint (\mathbf{j}\mathscr{E}_x-\mathbf{i}\mathscr{E}_y)\,\phi\,dx'\,dy'. \end{aligned}\right| \tag{3·2}$$

Equations (3·1), then give for $z>0$

$$\left.\begin{aligned} E_x &= \frac{\partial B_y}{\partial z}-ikA_x-\frac{\partial \Phi}{\partial x}, & H_x &= -\frac{\partial A_y}{\partial z}-ikB_x-\frac{\partial \Psi}{\partial x}, \\ E_y &= -\frac{\partial B_x}{\partial z}-ikA_y-\frac{\partial \Phi}{\partial y}, & H_y &= \frac{\partial A_x}{\partial z}-ikB_y-\frac{\partial \Psi}{\partial y}, \\ E_z &= -\frac{\partial \Phi}{\partial z}-\frac{\partial B_y}{\partial x}+\frac{\partial B_x}{\partial y}, & H_z &= -\frac{\partial \Psi}{\partial z}+\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}, \end{aligned}\right\} \tag{3·3}$$

but when $z<0$, the expressions on the right-hand sides of equations (3·3) are identically zero.

Now Φ, Ψ, A_x, A_y, B_x, B_y are all even functions of z, and so, when the sign of z is changed, the sign of all the derivatives with respect to z must be changed. Hence, when $z>0$, the first term on the right-hand side of each of equations (3·3) is equal to the sum of the second and third terms. It follows that the components of **E** and **H** can be expressed either in terms of the magnetic potentials or in terms of the electric potentials, the two sets of formulae being

$$\left.\begin{aligned} &E_x = 2\frac{\partial B_y}{\partial z}, && E_y = -2\frac{\partial B_x}{\partial z}, && E_z = -2\frac{\partial B_y}{\partial x}+2\frac{\partial B_x}{\partial y},\\ &H_x = -2ikB_x-2\frac{\partial \Psi}{\partial x}, && H_y = -2ikB_y-2\frac{\partial \Psi}{\partial y}, && H_z = -2\frac{\partial \Psi}{\partial z}, \end{aligned}\right\}.$$

and

$$\left.\begin{aligned} &E_x = -2ikA_x-2\frac{\partial \Phi}{\partial x}, && E_y = -2ikA_y-2\frac{\partial \Phi}{\partial y}, && E_z = -2\frac{\partial \Phi}{\partial z},\\ &H_x = -2\frac{\partial A_y}{\partial z}, && H_y = 2\frac{\partial A_x}{\partial z}, && H_z = 2\frac{\partial A_y}{\partial x}-2\frac{\partial A_x}{\partial y}, \end{aligned}\right\}$$

when $z>0$. But these are precisely equations (2·1) and (2·5), with $U=-2B_y$, $V=2B_x$, $W=2\Psi$, $U'=2A_y$, $V'=-2A_x$, $W'=2\Phi$. Moreover, the relations (2·3) and (2·7) are simply the relations $ik\Psi+\operatorname{div}\mathbf{B}=0$, $ik\Phi+\operatorname{div}\mathbf{A}=0$ connecting the scalar and vector potentials. The formulae for the half-space $z<0$ can be obtained in a similar way.

4. THE INTEGRAL EQUATIONS OF DIFFRACTION BY A PLANE SCREEN

THEOREM 4. *Let an electromagnetic field* $\mathbf{E}^i$, $\mathbf{H}^i$ *be incident in the half-space* $z>0$ *on a perfectly conducting screen in the plane* $z=0$, *the holes* in the screen being denoted by* S_1, *the metal of the screen by* S_2. *Then the total field in the half-space* $z<0$ *is*

$$\left.\begin{aligned} &E_x = \frac{\partial u}{\partial z}, && E_y = \frac{\partial v}{\partial z}, && E_z = -\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y},\\ &H_x = ikv+\frac{\partial w}{\partial x}, && H_y = -iku+\frac{\partial w}{\partial y}, && H_z = \frac{\partial w}{\partial z}, \end{aligned}\right\} \tag{4·11}$$

where

$$u=\frac{1}{2\pi}\iint_{S_1} e_x\phi\,dx'\,dy',\quad v=\frac{1}{2\pi}\iint_{S_1} e_y\phi\,dx'\,dy',\quad w=\frac{1}{2\pi}\iint_{S_1} h_z\phi\,dx'\,dy' \tag{4·12}$$

and

$$ikw=-\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}. \tag{4·13}$$

The functions e_x, e_y, h_z, *which are connected by the relation*

$$ikh_z=-\frac{\partial e_y}{\partial x'}+\frac{\partial e_x}{\partial y'}, \tag{4·14}$$

---

* If S_1 extends to infinity, the expressions $E^i_x e^{-ikr}$, $E^i_y e^{-ikr}$, $H^i_z e^{-ikr}$ with $z \doteq 0$ are assumed to tend to zero as $r=\sqrt{(x^2+y^2)}\to\infty$, uniformly in all directions on S_1.

satisfy the integral equations

$$\frac{\partial}{\partial x}\iint_{S_1} h_z\phi_0\,dx'\,dy' + ik\iint_{S_1} e_y\phi_0\,dx'\,dy' = 2\pi H_x^i(x,y,0), \tag{4·15}$$

$$\frac{\partial}{\partial y}\iint_{S_1} h_z\phi_0\,dx'\,dy' - ik\iint_{S_1} e_x\phi_0\,dx'\,dy' = 2\pi H_y^i(x,y,0), \tag{4·16}$$

$$\frac{\partial}{\partial x}\iint_{S_1} e_x\phi_0\,dx'\,dy' + \frac{\partial}{\partial y}\iint_{S_1} e_y\phi_0\,dx'\,dy' = -2\pi E_z^i(x,y,0), \tag{4·17}$$

when $(x, y, 0)$ is a point of S_1; ϕ_0 denotes the value of ϕ when $z = 0$. The expressions which would occur on the right-hand sides of the integral equations if $(x, y, 0)$ were a point of S_2, are unknown.

The total field everywhere will be

$$\mathbf{E} = \mathbf{E}^i + \mathbf{E}^s, \quad \mathbf{H} = \mathbf{H}^i + \mathbf{H}^s, \tag{4·2}$$

where $\mathbf{E}^s$, $\mathbf{H}^s$ is the scattered field, which has to be determined so that the tangential components of $\mathbf{E}^s$ and the normal component of $\mathbf{H}^s$ vanish on S_2. The boundary values of E_x^s, E_y^s, H_z^s are therefore

$$\mathscr{E}_x = -E_x^i + e_x, \quad \mathscr{E}_y = -E_y^i + e_y, \quad \mathscr{H}_z = -H_z^i + h_z,$$

where e_x, e_y, h_z vanish on S_2 but are unknown on S_1. Moreover, e_x, e_y, h_z are not independent, but satisfy (4·14) in virtue of Maxwell's equations.

Theorems 1 and 2 are valid since the boundary values are continuous except on the rim of the screen. The functions U, V, W of (2·2) become

$$U = -\frac{1}{2\pi}\iint E_x^i\phi\,dx'\,dy' + \frac{1}{2\pi}\iint_{S_1} e_x\phi\,dx'\,dy' = U^r + u,$$

$$V = -\frac{1}{2\pi}\iint E_y^i\phi\,dx'\,dy' + \frac{1}{2\pi}\iint_{S_1} e_y\phi\,dx'\,dy' = V^r + v,$$

$$W = -\frac{1}{2\pi}\iint H_z^i\phi\,dx'\,dy' + \frac{1}{2\pi}\iint_{S_1} h_z\phi\,dx'\,dy' = W^r + w,$$

in an obvious notation, where u, v, w must evidently satisfy (4·13). Hence, by (2·1) and (2·4),

$$E_x^s = \mp\frac{\partial U^r}{\partial z} \mp \frac{\partial u}{\partial z} = E_x^r \mp \frac{\partial u}{\partial z},$$

where the upper or lower sign is taken according as $z > 0$ or < 0. Therefore, by (4·2),

$$E_x = E_x^i + E_x^r \mp \frac{\partial u}{\partial z},$$

and similarly for the other components.

It is at this point that the fact that $\mathbf{E}^i$, $\mathbf{H}^i$ is incident on the screen in the half-space $z > 0$ comes into the argument. If there were no holes in the screen, the total field would be null in $z<0$, but would be $\mathbf{E}^0$, $\mathbf{H}^0$, say, in $z>0$. But U^r, V^r, W^r are independent of the form of the holes S_1, and u, v, w are identically zero when there are no holes. Hence

$$\mathbf{E}^i+\mathbf{E}^r = \mathbf{E}^0 \text{ or } 0, \quad \mathbf{H}^i+\mathbf{H}^r = \mathbf{H}^0 \text{ or } 0,$$

according as $z>0$ or <0. In fact, $\mathbf{E}^r$, $\mathbf{H}^r$ is the ordinary reflected field when $\mathbf{E}^i$, $\mathbf{H}^i$ is incident in $z>0$ on a perfectly reflecting plane $z=0$. It has thus been proved that the total field (4·2) is given by

$$E_x = E_x^0-\frac{\partial u}{\partial z}, \qquad E_y = E_y^0-\frac{\partial v}{\partial z}, \qquad E_z = E_z^0+\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y},$$

$$H_x = H_x^0-ikv-\frac{\partial w}{\partial x}, \quad H_y = H_y^0+iku-\frac{\partial w}{\partial y}, \quad H_z = H_z^0-\frac{\partial w}{\partial z},$$

when $z>0$, and by

$$E_x = \frac{\partial u}{\partial z}, \qquad E_y = \frac{\partial v}{\partial z}, \qquad E_z = -\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y},$$

$$H_x = ikv+\frac{\partial w}{\partial x}, \quad H_y = -iku+\frac{\partial w}{\partial y}, \quad H_z = \frac{\partial w}{\partial z},$$

when $z<0$. The equations of the latter set are equations (4·11) of theorem 4.

By their mode of construction, E_x, E_y, H_z are continuous across the holes S_1 of the screen. e_x, e_y, h_z must be chosen so that the remaining components are also continuous across S_1. This requires that

$$\frac{\partial w}{\partial x}+ikv = \tfrac{1}{2}H_x^0, \quad \frac{\partial w}{\partial y}-iku = \tfrac{1}{2}H_y^0, \quad \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y} = \tfrac{1}{2}E_z^0, \tag{4·3}$$

when $z=0$ and $(z,x,0)$ is a point of S_1.

Now it has already been seen that H_x^r, H_y^r, E_z^r are odd functions of z, and also that $\mathbf{E}^r = -\mathbf{E}^i$, $\mathbf{H}^r = -\mathbf{H}^i$ when $z<0$. Hence, as $z\to+0$,

$$H_x^0\to 2H_x^i, \quad H_y^0\to 2H_y^i, \quad E_z^0\to 2E_z^i.$$

Therefore equations (4·3) may be written as

$$\frac{\partial w}{\partial x}+ikv = H_x^i, \quad \frac{\partial w}{\partial y}-iku = H_y^i, \quad \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y} = -E_z^i$$

when $z=0$ and $(x,y,0)$ is a point of S_1; and these are the integral equations (4·15), (4·16) and (4·17) of theorem 4.

The expressions on the right-hand sides of the integral equations are unknown on S_2. For example

$$\frac{\partial w}{\partial x}+iku = H_x(x,y,-0)$$

when $z=0$ and $(x,y,0)$ is a point of S_2; and similarly for the other equations. The integral equations are therefore essentially of the form

$$\iint_{S_1} f(x',y')\,K(x,y;x',y')\,dx'\,dy' = F(x,y),$$

where $F(x,y)$ is given on S_1 but unknown on S_2. The radiation condition does, however, give some information about the asymptotic behaviour of $F(x,y)$ on S_2 when S_2 extends to infinity, and this may be of importance in a rigorous solution.

The three integral equations are not independent, since (4·17) is a consequence of (4·15) and (4·16). It is, nevertheless, useful to retain all three with a view to finding approximate solutions. Theorem 4 is very suitable for dealing with diffraction by small holes in a plane screen, but would be of little use in the case of a small obstacle. The complementary formulation given below in theorem 5 is suitable for small obstacle problems.

THEOREM 5. *Let an electromagnetic field* $\mathbf{E}^i$, $\mathbf{H}^i$ *be incident in the half-space* $z > 0$ *on a perfectly conducting screen in the plane* $z = 0$, *the holes in the screen being denoted by* S_1, *the metal* of the screen by* S_2. *Then the total field everywhere is*

$$\left.\begin{aligned} E_x &= E^i_x + ikv' - \frac{\partial w'}{\partial x}, & E_y &= E^i_y - iku' - \frac{\partial w'}{\partial y}, & E_z &= E^i_z - \frac{\partial w'}{\partial z}, \\ H_x &= H^i_x - \frac{\partial u'}{\partial z}, & H_y &= H^i_y - \frac{\partial v'}{\partial z}, & H_z &= H^i_z + \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}, \end{aligned}\right\} \tag{4·41}$$

where

$$u' = \frac{1}{2\pi}\iint_{S_2} h_x\phi\, dx'\, dy', \quad v' = \frac{1}{2\pi}\iint_{S_2} h_y\phi\, dx'\, dy', \quad w' = \frac{1}{2\pi}\iint_{S_2} e_z\phi\, dx'\, dy' \tag{4·42}$$

and

$$ikw' = \frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y}. \tag{4·43}$$

The functions h_x, h_y, e_z, *which are connected by the relation*

$$ike_z = \frac{\partial h_y}{\partial x'} - \frac{\partial h_x}{\partial y'}, \tag{4·44}$$

satisfy the integral equations

$$\frac{\partial}{\partial x}\iint_{S_2} e_z\phi_0\, dx'\, dy' - ik\iint_{S_2} h_y\phi_0\, dx'\, dy' = 2\pi E^i_x(x,y,0), \tag{4·45}$$

$$\frac{\partial}{\partial y}\iint_{S_2} e_z\phi_0\, dx'\, dy' + ik\iint_{S_2} h_x\phi_0\, dx'\, dy' = 2\pi E^i_y(x,y,0), \tag{4·46}$$

$$\frac{\partial}{\partial x}\iint_{S_2} h_x\phi_0\, dx'\, dy' + \frac{\partial}{\partial y}\iint_{S_2} h_y\phi_0\, dx'\, dy' = -2\pi H^i_z(x,y,0), \tag{4·47}$$

when $(z,y,0)$ *is a point of* S_2; ϕ_0 *denotes the value of* ϕ *when* $x = 0$. *The expressions which would occur on the right-hand sides of the integral equations if* $(x,y,0)$ *were a point of* S_1 *are unknown.*

Using the notation of the proof of theorem 4, it is known that H^s_x, H^s_y, E^s_z are odd functions of z, and so

$$H^s_x \to \pm h_x, \quad H^s_y \to \pm h_y, \quad E^s_z \to \pm e_z,$$

---

* If S_2 extends to infinity, the expressions $H^i_x e^{-ikr}$, $H^i_y e^{-ikr}$, $E^i_z e^{-ikr}$ with $z = 0$ are assumed to tend to zero as $r = \sqrt{(x^2+y^2)} \to \infty$, uniformly in all directions on S_2.

as $z \to \pm 0$. But since H^s_x, H^s_y, E^s_z are continuous across the holes in the screen, h_x, h_y and e_z vanish on S_1. It follows from theorem 3 that, if u', v', w' are defined by (4·42), then

$$E^s_x = ikv' - \frac{\partial w'}{\partial x}, \quad E^s_y = -iku' - \frac{\partial w'}{\partial y}, \quad E^s_z = -\frac{\partial w'}{\partial z},$$

$$H^s_x = -\frac{\partial u'}{\partial y}, \qquad H^s_y = -\frac{\partial v'}{\partial z}, \qquad H^s_z = \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y},$$

everywhere. The total field $\mathbf{E} = \mathbf{E}^i + \mathbf{E}^s$, $\mathbf{H} = \mathbf{H}^i + \mathbf{H}^s$ is therefore given by (4·41) everywhere. Since, however, E_x, E_y, H_z vanish on S_2, then

$$\frac{\partial w'}{\partial x} - ikv' = E^i_x, \quad \frac{\partial w'}{\partial y} + iku' = E^i_y, \quad \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} = -H^i_z,$$

when $z = 0$ and $(x, y, 0)$ is a point of S_2; and these are the integral equations (4·45), (4·46) and (4·47) of theorem 5.

The expressions which would occur on the right-hand sides of the integral equations if $(x, y, 0)$ were a point of S_1 are unknown; for example,

$$\frac{\partial w'}{\partial x} - ikv' = -E^s_x(x, y, 0),$$

and E^s_x is unknown on S_1.

The three integral equations are not independent, since (4·47) is a consequence of (4·45) and (4·46). Again all three are retained with a view to finding approximate solutions.

5. SOMMERFELD'S DIFFRACTION PROBLEM

Theorems 4 and 5 will now be applied to Sommerfeld's problem of the diffraction of plane-polarized waves by a perfectly conducting half-plane. The axes are chosen so that the screen lies in the plane $z = 0$ and so that the waves are incident on it in $z > 0$; the edge of the screen is taken to be the axis of y so that the hole S_1 is $z = 0$, $x > 0$, and the metal of the screen S_2 is $z = 0, x < 0$. There are two cases to be considered.

In the first case, the plane of polarization is parallel to the edge of the screen, the incident field being

$$\mathbf{E}^i = (-\mathbf{i}\sin\alpha + \mathbf{k}\cos\alpha)\, e^{ikx\cos\alpha + ikz\sin\alpha}, \quad \mathbf{H}^i = \mathbf{j}e^{ikx\cos\alpha + ikz\sin\alpha},$$

where $0 < \alpha < \pi$. The problem is a two-dimensional one; in the total field $\mathbf{E}$ will be everywhere perpendicular to the axis of y, $\mathbf{H}$ everywhere parallel to the axis of y, and both vectors will be independent of y. Hence the functions e_x, e_y, h_z of theorem 4 will not involve y', and also e_y, and therefore h_z, will be identically zero. There is, then, only the function e_x to be found.

The integral equation (4·15) of theorem 4 will be satisfied identically, and (4·17) will be an immediate consequence of (4·16). There is then only one integral equation to solve, namely,

$$-ik \iint_{S_1} e_x \frac{e^{-ik\rho}}{\rho}\, dx'\, dy' = 2\pi e^{ikx\cos\alpha},$$

when $(x, y, 0)$ is a point of S_1; here $\rho^2 = (x-x')^2+(y-y')^2$. As e_x does not involve y', one can integrate with respect to y', to obtain

$$\int_0^\infty e_x(x')\, H_0^{(2)}(k\,|\,x-x'\,|)\,dx' = -\frac{2}{k}\,e^{ikx\cos\alpha}, \tag{5·1}$$

when $x>0$. When (5·1) has been solved for $e_x(x')$, the total field behind the screen will be given by equations (4·11).

It is well known that this case of Sommerfeld's problem is mathematically identical with an acoustical problem in which plane waves with velocity potential

$$e^{ikx\cos\alpha+ikz\sin\alpha+ikct}$$

are incident on the perfectly reflecting screen $z=0$, $x<0$; the velocity potential satisfies the same differential equation and the same boundary conditions as the component H_y of the magnetic vector. I have derived elsewhere (1945) the integral equation (5·1) for this acoustical problem, and have solved it by means of complex Fourier integrals. The solution leads to Sommerfeld's formulae in both the acoustical and electromagnetic problems.

In the second case, the plane of polarization is perpendicular to the edge of the screen, the incident field being

$$\mathbf{E}^i = \mathbf{j}e^{ikx\cos\alpha+ikz\sin\alpha}, \quad \mathbf{H}^i = (\mathbf{i}\sin\alpha-\mathbf{k}\cos\alpha)\,e^{ikx\cos\alpha+ikz\sin\alpha},$$

where $0<\alpha<\pi$. The problem is again a two-dimensional one, with **E** and **H** independent of y and **E** parallel to the axis of y. The function h_x of theorem 5 is independent of y' and h_y and e_z are identically zero.

The integral equation (4·45) of theorem 5 will be satisfied identically, and (4·47) will be an immediate consequence of (4·46). Thus, again, there is only one integral equation to solve, namely,

$$ik\iint_{S_2} h_x(x')\,\frac{e^{-ik\rho}}{\rho}\,dx'\,dy' = 2\pi e^{ikx\cos\alpha},$$

when $(x, y, 0)$ is a point of S_2. After integration with respect to y', this becomes

$$\int_{-\infty}^0 h_x(x')\, H_0^{(2)}(k\,|\,x-x'\,|)\,dx' = \frac{2}{k}\,e^{ikx\cos\alpha}, \tag{5·2}$$

when $x<0$. When (5·2) has been solved for $h_x(x')$, the total field everywhere is given by (4·41).

It will be seen that the integral equations (5·1) and (5·2) are almost identical. It will be seen later that each implies the other, in virtue of the rigorous form of Babinet's principle to be proved in § 7. The equation (5·2) was first obtained by Magnus (1941), who solved it by rather difficult analysis involving series of Bessel functions and verified that the solution led to Sommerfeld's formula in the case of normal incidence. The Fourier integral solution is much simpler and more direct.

6. DIFFRACTION BY A SMALL CIRCULAR HOLE AND A SMALL CIRCULAR DISK

These problems were first solved by Rayleigh in 1897 on the assumption that the radius of the hole or disk is small compared with the wave-length, by his method of regarding the field near the hole or disk as quasi-static. The hole problem has recently been solved again by Bethe (1944), who regarded the field behind the screen as being due to a fictitious magnetic charge and current sheet in the hole. Both methods lead to the same result, which differs surprisingly from the classical results for the case when the radius of the hole is large compared with the wave-length. It will now be shown that the methods developed in the present paper lead very simply to the solution of both problems.

In the first problem, suppose that the field $\mathbf{E}^i$, $\mathbf{H}^i$ is incident in $z>0$ on the perfectly conducting plane $z=0$, which is pierced by a small hole S_1, on which $x^2+y^2 \leqslant a^2$, where $ka \ll 1$. Theorem 4 gives three integral equations, which will be solved approximately, making, as Bethe does, the following assumptions:

(i) since $ka \ll 1$, H_x^i, H_y^i, E_z^i may be regarded as constant over S_1;

(ii) since $\phi_0 = e^{-ik\rho}/\rho$, where $\rho^2 = (x-x')^2+(y-y')^2$, and since $k\rho$ does not exceed $2ka$, where (x,y) and (x',y') are points of S_1, ϕ_0 is replaced by $1/\rho$. The integral equations then become

$$\frac{\partial}{\partial x}\iint_{S_1} h_z \frac{dx'\,dy'}{\rho} + ik\iint_{S_1} e_y \frac{dx'\,dy'}{\rho} = 2\pi \bar{H}_x^i, \tag{6·1}$$

$$\frac{\partial}{\partial y}\iint_{S_1} h_z \frac{dx'\,dy'}{\rho} - ik\iint_{S_1} e_x \frac{dx'\,dy'}{\rho} = 2\pi \bar{H}_y^i, \tag{6·2}$$

$$\frac{\partial}{\partial x}\iint_{S_1} e_x \frac{dx'\,dy'}{\rho} + \frac{\partial}{\partial y}\iint_{S_1} e_y \frac{dx'\,dy'}{\rho} = -2\pi \bar{E}_z^i, \tag{6·3}$$

when (x,y) is a point of S_1; here $\bar{H}_x^i$ denotes the value of H_x^i at O, and similarly in the other equations.

The three integral equations are evidently not independent, nor are the three unknown functions, since

$$ikh_z = -\frac{\partial e_y}{\partial x} + \frac{\partial e_x}{\partial y};$$

but it is desirable to keep all three equations in an approximate solution. The equations will be solved by using the formulae*

$$\iint_{S_1} \frac{\sqrt{(a^2-r'^2)}}{\rho}\,dx'\,dy' = \tfrac{1}{2}\pi^2(a^2-\tfrac{1}{2}r^2),$$

$$\iint_{S_1} \frac{x'}{\sqrt{(a^2-r'^2)}}\frac{dx'\,dy'}{\rho} = \tfrac{1}{2}\pi^2 x,$$

$$\iint_{S_1} \frac{y'}{\sqrt{(a^2-r'^2)}}\frac{dx'\,dy'}{\rho} = \tfrac{1}{2}\pi^2 y,$$

---

* The first formula is suggested by the expression for the potential of a uniform solid ellipsoid at an internal point and can be proved by direct integration using polar co-ordinates with (x, y) as pole. The second and third follow from the first. The analysis is suggested by that of Bethe.

when (x, y) is a point of S_1; for brevity r^2 and r'^2 have been written instead of x^2+y^2 and $x'^2+y'^2$. These formulae and (6·3) suggest that a suitable solution would be

$$e_x = \frac{Ax'}{\sqrt{(a^2-r'^2)}}, \quad e_y = \frac{A'y'}{\sqrt{(a^2-r'^2)}}.$$

But this would give $$ikh_z = -\frac{\partial e_y}{\partial x} + \frac{\partial e_x}{\partial y} = \frac{(A-A')\,x'y'}{(a^2-r'^2)^{\frac{3}{2}}},$$

and the integral for w would diverge unless $A = A'$, when w would be identically zero. Again (6·1) and (6·2) suggest that

$$ikh_z = \frac{Bx'}{\sqrt{(a^2-r'^2)}} + \frac{Cy'}{\sqrt{(a^2-r'^2)}}.$$

Hence the trial solution may be given as

$$e_x = \frac{Ax'}{\sqrt{(a^2-r'^2)}} - C\sqrt{(a^2-r'^2)},$$

$$e_y = \frac{Ay'}{\sqrt{(a^2-r'^2)}} + B\sqrt{(a^2-r'^2)},$$

$$ikh_z = \frac{Bx'}{\sqrt{(a^2-r'^2)}} + \frac{Cy'}{\sqrt{(a^2-r'^2)}},$$

which satisfy (4·14). Moreover, although these functions become infinite on the rim of S_1, the integrals u, v, w converge and satisfy (4·13).

When the trial solution is substituted in the integral equations, then

$$B\{1-k^2(a^2-\tfrac{1}{2}r^2)\} = \frac{4ik}{\pi}\bar{H}^i_x + Ak^2y,$$

$$C\{1-k^2(a^2-\tfrac{1}{2}r^2)\} = \frac{4ik}{\pi}\bar{H}^i_y - Ak^2x,$$

$$A = -\frac{\pi}{2}\bar{E}^i_z + \tfrac{1}{2}(By - Cx),$$

and hence $$A = -\frac{2}{\pi}\bar{E}^i_z + O(ka),$$

$$B = \left\{\frac{4ik}{\pi}\bar{H}^i_x + Ak^2y\right\}\{1+O(k^2a^2)\},$$

$$C = \left\{\frac{4ik}{\pi}\bar{H}^i_y - Ak^2x\right\}\{1+O(k^2a^2)\}.$$

Therefore, neglecting a multiplier $\{1+O(ka)\}$, it follows that

$$A = -\frac{2}{\pi}\bar{E}^i_z, \quad B = \frac{4ik}{\pi}\bar{H}^i_x, \quad C = \frac{4ik}{\pi}\bar{H}^i_y,$$

and so the approximate solution of the set of integral equations is

$$e_x = -\frac{2}{\pi}\bar{E}^i_z \frac{x'}{\sqrt{(a^2-r'^2)}} - \frac{4ik}{\pi}\bar{H}^i_y \sqrt{(a^2-r'^2)}, \tag{6·4}$$

$$e_y = -\frac{2}{\pi}\bar{E}^i_z \frac{y'}{\sqrt{(a^2-r'^2)}} + \frac{4ik}{\pi}\bar{H}^i_x \sqrt{(a^2-r'^2)}, \tag{6·5}$$

$$h_z = \frac{4}{\pi}\bar{H}^i_x \frac{x'}{\sqrt{(a^2-r'^2)}} + \frac{4}{\pi}\bar{H}^i_y \frac{y'}{\sqrt{(a^2-r'^2)}}. \tag{6·6}$$

These are Bethe's equations (33) and (33*a*) in a different notation. The field behind the screen can now be found by calculating u, v, w and substituting in (4·11).

To find the field when the distance $r = \sqrt{(x^2+y^2+z^2)}$ is large compared with the wave-length, the approximation

$$\frac{e^{-ikR}}{R} = \frac{e^{-ikr}}{r}\left\{1 + ik\frac{xx'+yy'}{r}\right\},$$

can be used, which is valid since kx' and ky' are small on S_1, and hence

$$u = -\frac{2ika^3}{3\pi}\left(\frac{x}{r}\bar{E}^i_z + 2\bar{H}^i_y\right)\frac{e^{-ikr}}{r},$$

$$v = -\frac{2ika^3}{3\pi}\left(\frac{y}{r}\bar{E}^i_z - 2\bar{H}^i_x\right)\frac{e^{-ikr}}{r},$$

$$w = \frac{4ika^3}{3\pi}\left(\frac{x}{r}\bar{H}^i_x + \frac{y}{r}\bar{H}^i_y\right)\frac{e^{-ikr}}{r}.$$

A straightforward calculation shows that the total field at great distances behind the screen is

$$E_x = -\frac{2k^2a^3}{3\pi}\left(\frac{zx}{r^2}\bar{E}^i_z + \frac{2z}{r}\bar{H}^i_y\right)\frac{e^{-ikr}}{r},$$

$$E_y = -\frac{2k^2a^3}{3\pi}\left(\frac{yz}{r^2}\bar{E}^i_z - \frac{2z}{r}\bar{H}^i_x\right)\frac{e^{-ikr}}{r},$$

$$E_z = \frac{2k^2a^3}{3\pi}\left(\frac{x^2+y^2}{r^2}\bar{E}^i_z + \frac{2x}{r}\bar{H}^i_y - \frac{2y}{r}\bar{H}^i_x\right)\frac{e^{-ikr}}{r},$$

$$H_x = \frac{2k^2a^3}{3\pi}\left(\frac{y}{r}\bar{E}^i_z - 2\frac{y^2+z^2}{r^2}\bar{H}^i_x + \frac{2xy}{r^2}\bar{H}^i_y\right)\frac{e^{-ikr}}{r},$$

$$H_y = \frac{2k^2a^3}{3\pi}\left(-\frac{x}{r}\bar{E}^i_z + \frac{2xy}{r^2}\bar{H}^i_x - 2\frac{z^2+x^2}{r^2}\bar{H}^i_y\right)\frac{e^{-ikr}}{r},$$

$$H_z = \frac{4k^2a^3}{3\pi}\left(\frac{zx}{r^2}\bar{H}^i_x + \frac{yz}{r^2}\bar{H}^i_y\right)\frac{e^{-ikr}}{r}.$$

In spherical polar coordinates

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta,$$

where $\frac{1}{2}\pi < \theta < \pi$, these results take the simpler form

$$E_r = H_r = 0,$$

$$E_\theta = H_\phi = \frac{2k^2a^3}{3\pi}(-\bar{E}^i_z \sin\theta + 2\bar{H}^i_y \sin\phi - 2\bar{H}^i_y \cos\phi)\frac{e^{-ikr}}{r},$$

$$E_\phi = -H_\theta = \frac{4k^2a^3}{3\pi}(\bar{H}^i_x \cos\phi + \bar{H}^i_y \sin\phi)\cos\theta\frac{e^{-ikr}}{r},$$

which exhibit the transversality of the waves and the orthogonality of **E** and **H**.

These expressions for the total field agree with the formulae of Rayleigh (for plane incident waves) and Bethe.

To the complementary problem of diffraction by a small circular disk, theorem 5 is applied. Let the field $\mathbf{E}^i$, $\mathbf{H}^i$ be incident in $z > 0$ on a small circular disk S_2 in the plane $z = 0$; on S_2, $x^2 + y^2 \leqslant a^2$, where $ka \ll 1$. The integral equations will be solved approximately, making the following assumptions:

(i) E^i_x, E^i_y, H^i_z are regarded as constant over S_2;

(ii) ϕ_0 is replaced by $1/\rho$.

The integral equations then become

$$\frac{\partial}{\partial x}\iint_{S_2} e_z \frac{dx'\,dy'}{\rho} - ik\iint_{S_2} h_y \frac{dx'\,dy'}{\rho} = 2\pi\bar{E}^i_x,$$

$$\frac{\partial}{\partial y}\iint_{S_2} e_z \frac{dx'\,dy'}{\rho} + ik\iint_{S_2} h_x \frac{dx'\,dy'}{\rho} = 2\pi\bar{E}^i_y,$$

$$\frac{\partial}{\partial x}\iint_{S_2} h_x \frac{dx'\,dy'}{\rho} + \frac{\partial}{\partial y}\iint_{S_2} h_y \frac{dx'\,dy'}{\rho} = -2\pi\bar{H}^i_z,$$

when (x, y) is a point of S_2; here $\bar{E}^i_x$ denotes the value of E^i_x at O, and similarly in the other equations.

Apart from differences in notation, these are precisely the integral equations (6·1), (6·2) and (6·3) of the small circular hole problem, and the approximate solution

$$h_x = -\frac{2}{\pi}\bar{H}^i_z \frac{x'}{\sqrt{(a^2 - r'^2)}} + \frac{4ik}{\pi}\bar{E}^i_y \sqrt{(a^2 - r'^2)},$$

$$h_y = -\frac{2}{\pi}\bar{H}^i_z \frac{y'}{\sqrt{(a^2 - r'^2)}} - \frac{4ik}{\pi}\bar{E}^i_x \sqrt{(a^2 - r'^2)},$$

$$e_z = \frac{4}{\pi}\bar{E}^i_x \frac{x'}{\sqrt{(a^2 - r'^2)}} + \frac{4}{\pi}\bar{E}^i_y \frac{y'}{\sqrt{(a^2 - r'^2)}},$$

can be written down at once from (6·4), (6·5) and (6·6). It readily follows that, when $kr \gg 1$,

$$u' = -\frac{2ika^3}{3\pi}\left(\frac{x}{r}\bar{H}^i_z - 2\bar{E}^i_y\right)\frac{e^{-ikr}}{r},$$

$$v' = -\frac{2ika^3}{3\pi}\left(\frac{y}{r}\bar{H}^i_y + 2\bar{E}^i_x\right)\frac{e^{-ikr}}{r},$$

$$w' = \frac{4ika^3}{3\pi}\left(\frac{x}{r}\bar{E}^i_h + \frac{y}{r}\bar{E}^i_y\right)\frac{e^{-ikr}}{r},$$

and hence that the scattered field everywhere, in spherical polar co-ordinates, is

$$E^s_r = H^s_r = 0,$$

$$E^s_\theta = H^s_\phi = \frac{4k^2a^3}{3\pi}(\bar{E}^i_x \cos\phi + \bar{E}^i_y \sin\phi)\cos\theta \frac{e^{-ikr}}{r},$$

$$E^s_\phi = -H^s_\theta = -\frac{2k^2a^3}{3\pi}(\bar{H}^i_z \sin\theta + 2\bar{E}^i_x \sin\phi - 2\bar{E}^i_y \cos\phi)\frac{e^{-ikr}}{r}.$$

These expressions for the scattered field agree with the formulae of Rayleigh for plane incident waves.

The methods of the present paper can also be applied to the problem of diffraction by a narrow slit or a narrow blade when the incident field is the same in all planes perpendicular to the edges of the slit or blade. As no new mathematical ideas are involved and as the results agree with Rayleigh's, the discussion of these problems is omitted.

7. BABINET'S PRINCIPLE

When electromagnetic waves fall on complementary plane screens, complementary diffraction patterns are formed. Babinet's principle in optics asserts that the sum of the electric force at any point in one diffraction pattern and the electric force at the same point in the complementary pattern is equal to the undisturbed electric force at that point when no screen is present. This principle is usually based on the classical Kirchhoff theory of diffraction by a black screen, a somewhat insecure foundation.

An examination of the formulae for diffraction by a small circular hole and a small circular disk given in § 6 shows that Babinet's principle in the usually accepted sense is not even approximately true for perfectly reflecting screens; this is also the case for diffraction by a narrow slit and a narrow blade. This was noticed in 1941 by H. G. Booker,* who observed that the principle would still hold for perfectly reflecting screens in the cases discussed by Rayleigh provided that, in passing from a screen to its complement, one made also the transformation $\mathbf{E} \to \mathbf{H}$, $\mathbf{H} \to -\mathbf{E}$; and he was able to prove the principle in this form by considering wave-motions in a two-sheeted space. It will now be shown that Babinet's principle in its rigorous form is actually an immediate and simple consequence of the complementary theorems 4 and 5 of the present paper.

THEOREM 6. (i) *Let the electromagnetic field*

$$\mathbf{E}^i = \mathbf{F}, \quad \mathbf{H}^i = \mathbf{G}$$

be incident in $z > 0$ *on a perfectly conducting screen in the plane* $z = 0$*; the holes in the screen are denoted by* Σ*, the metal of the screen by* Σ'*. Let the total field in* $z < 0$ *be* $\mathbf{E}^1, \mathbf{H}^1$.

(ii) *Let the complementary electromagnetic field*

$$\mathbf{E}^i = -\mathbf{G}, \quad \mathbf{H}^i = \mathbf{F}$$

---

* I wish to express my gratitude to Dr Booker for drawing my attention to this interesting problem and for giving me permission to refer here to his unpublished work.

be incident in $z>0$ on the complementary perfectly conducting screen in the plane $z=0$; the holes in the screen are Σ', the metal of the screen Σ. Let the total field in $z<0$ be $\mathbf{E}^2$, $\mathbf{H}^2$.

Then Babinet's principle in its rigorous form asserts that

$$\mathbf{E}^1+\mathbf{H}^2=\mathbf{F}, \quad \mathbf{H}^1-\mathbf{E}^2=\mathbf{G}.$$

Theorem 4 is first applied to problem (i). By equations (4·11), the total field $\mathbf{E}^1$, $\mathbf{H}^1$ in $z<0$ is

$$\left.\begin{aligned} E^1_x &= \frac{\partial u}{\partial z}, & E^1_y &= \frac{\partial v}{\partial z}, & E^1_z &= -\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}, \\ H^1_x &= ikv+\frac{\partial w}{\partial x}, & H^1_y &= -iku+\frac{\partial w}{\partial y}, & H^1_z &= \frac{\partial w}{\partial z}, \end{aligned}\right\} \tag{7·1}$$

where

$$u=\frac{1}{2\pi}\iint_\Sigma f\phi\,dx'\,dy', \quad v=\frac{1}{2\pi}\iint_\Sigma g\phi\,dx'\,dy', \quad w=\frac{1}{2\pi}\iint_\Sigma h\phi\,dx'\,dy'.$$

The functions f, g, h are the functions e_x, e_y, h_z of theorem 4; they are connected by

$$ikh=-\frac{\partial g}{\partial x'}+\frac{\partial f}{\partial y'},$$

and satisfy the integral equations

$$\left.\begin{aligned} \frac{\partial}{\partial x}\iint_\Sigma h\phi_0\,dx'\,dy'+ik\iint_\Sigma g\phi_0\,dx'\,dy' &= 2\pi G_x(x,y,0), \\ \frac{\partial}{\partial y}\iint_\Sigma h\phi_0\,dx'\,dy'-ik\iint_\Sigma f\phi_0\,dx'\,dy' &= 2\pi G_y(x,y,0), \\ \frac{\partial}{\partial x}\iint_\Sigma f\phi_0\,dx'\,dy'+\frac{\partial}{\partial y}\iint_\Sigma g\phi_0\,dx'\,dy' &= -2\pi F_z(x,y,0), \end{aligned}\right\} \tag{7·2}$$

when (x,y) is a point of Σ.

Next theorem 5 is applied to problem (ii). By equations (4·41), the total field everywhere is

$$\left.\begin{aligned} E^2_x &= -G_x+ikv'-\frac{\partial w'}{\partial x}, & E^2_y &= -G_y-iku'-\frac{\partial w'}{\partial y}, & E^2_z &= -G_z-\frac{\partial w'}{\partial z}, \\ H^2_x &= F_x-\frac{\partial u'}{\partial z}, & H^2_y &= F_y-\frac{\partial v'}{\partial z}, & H^2_z &= F_z+\frac{\partial u'}{\partial x}+\frac{\partial v'}{\partial y}, \end{aligned}\right\} \tag{7·3}$$

where

$$u'=\frac{1}{2\pi}\iint_\Sigma f'\phi\,dx'\,dy', \quad v'=\frac{1}{2\pi}\iint_\Sigma g'\phi\,dx'\,dy', \quad w'=\frac{1}{2\pi}\iint_\Sigma h'\phi\,dx'\,dy'.$$

The functions f', g', h' are the functions h_x, h_y, e_z of theorem 5; they are connected by

$$ikh'=\frac{\partial g'}{\partial x'}-\frac{\partial f'}{\partial y'},$$

and satisfy the integral equations

$$\left.\begin{aligned}\frac{\partial}{\partial x}\iint_{\Sigma} h'\phi_0\,dx'\,dy' - ik\iint_{\Sigma} g'\phi_0\,dx'\,dy' &= -2\pi G_x(x,y,0),\\ \frac{\partial}{\partial y}\iint_{\Sigma} h'\phi_0\,dx'\,dy' + ik\iint_{\Sigma} f'\phi_0\,dx'\,dy' &= -2\pi G_y(x,y,0),\\ \frac{\partial}{\partial x}\iint_{\Sigma} f'\phi_0\,dx'\,dy' + \frac{\partial}{\partial y}\iint_{\Sigma} g'\phi_0\,dx'\,dy' &= -2\pi F_z(x,y,0),\end{aligned}\right\} \qquad (7\cdot4)$$

when (x, y) is a point of Σ.

Comparing (7·2) and (7·4), it is seen that

$$f' = f, \quad g' = g, \quad h' = -h,$$

and hence that

$$u' = u, \quad v' = v, \quad w' = -w.$$

Equations (7·3) then become

$$\left.\begin{aligned}E^2_x &= -G_x + ikv + \frac{\partial w}{\partial x}, & E^2_y &= -G_y - iku + \frac{\partial w}{\partial y}, & E^2_z &= -G_z + \frac{\partial w}{\partial z},\\ H^2_x &= F_x - \frac{\partial u}{\partial z}, & H^2_y &= F_y - \frac{\partial v}{\partial z}, & H^2_z &= F_z + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}.\end{aligned}\right\} \qquad (7\cdot5)$$

From (7·1) and (7·5) the equations

$$\mathbf{E}^1 + \mathbf{H}^2 = \mathbf{F}, \quad \mathbf{H}^1 - \mathbf{E}^2 = \mathbf{G}$$

follow. This completes the proof of Babinet's principle.

REFERENCES

Baker & Copson 1939 *Huygens' Principle*, 102–109. Oxford.
Bethe 1944 *Phys. Rev.* (2) **66**, 163–182.
Copson 1945 *Quart. J. Math.* (in the Press).
Magnus 1941 *Z. Phys.* **117**, 168–179.
Rayleigh 1896 *Theory of Sound*, **2**, 107 and 109. London.
Rayleigh 1897 *Phil. Mag.* (5) **43**, 259–272; **44**, 28–52.
Schwarzschild 1902 *Math. Ann.* **55**, 177–247.

On the divergence difficulty of quantized field theories and the rigorous treatment of radiation reaction

By H. W. Peng, *University of Edinburgh*

(*Communicated by M. Born, F.R.S.—Received* 5 *April* 1945)

By an orthodox application of the perturbation theory to the general case of a quantized field, it is shown that the divergence difficulty hitherto encountered arises from a faulty application of the expansion method. The difficulty, in certain cases, disappears if the degeneracy of the unperturbed system is properly treated by the method of secular perturbation. Physically, this amounts to a rigorous treatment of the radiation reaction, in such cases where its effect is strong.

The existence of elementary particles (photons, electrons, etc.) in integral numbers has been successfully explained by the application of quantum mechanics for the treatment of fields (Maxwell's field, Dirac (1927), and Jordan & Pauli (1928); Dirac's field, Jordan & Wigner (1928); etc.). The interaction of the elementary particles is usually introduced by coupling the corresponding fields, for which a general description will be given in § 1 below. The effects (self-energies and cross-sections) due to the interaction have been hitherto treated by the expansion method—a power-series expansion with respect to the interaction constants. All terms of the resultant expansions for the self-energies and the cross-sections, excepting the term of the lowest degree in each expansion, are found to diverge. This difficulty has led many physicists to doubt either the usual formulation of the field theory or the usual procedure of quantization (which is generally known as that of Heisenberg & Pauli (1929)). But it will be shown in § 2 below that for *any* quantized field, the system is degenerate in the sense of the perturbation theory, and hence the expansion method must give way to the method of secular perturbation. Therefore no premature departure from the usual formulation of the quantized field theories will be considered in this paper. On the other hand, the mathematical treatment of secular perturbation which will be developed in §§ 3–5 (or some further improvement of it) will have to be used for the derivation of the self-energies and cross-sections in any new formulation of the quantized field theories.

In § 3, the usual method of secular perturbation for degenerate systems will be presented in a form which can easily be adapted to the case of continuous spectra. The general solution of the secular problem will be given in §§ 4–5, where it will be seen that most of the stationary states of the secular problem describe collisions of the elementary particles. For the calculation of the cross-sections for the collisions one has to solve some integral equations which differ from those proposed by Heitler (1941) and Wilson (1941), in their theory of radiation damping, merely by not neglecting or omitting the self-energies (their treatment is only preliminary and is integrated in the present treatment). The present treatment, which in fact is just the orthodox perturbation treatment and can be continued to any desired order of approximation, can therefore be described physically as a *rigorous* treatment of the radiation reaction. The self-energies are also obtained from the stationary states of the secular problem, and are influenced by the radiation reaction. As will be demonstrated in § 6 with the help of a simple example, in which the radiation reaction is strong, the present treatment leads to no divergence difficulty where the usual treatment does.

1. GENERAL DESCRIPTION OF THE INTERACTION OF ELEMENTARY PARTICLES BY MEANS OF A QUANTIZED FIELD

All elementary particles can be shown to arise from the corresponding field theories by a definite procedure called 'quantization' (cf. (Dirac (1935), chap. XIII; Pauli (1941), part II). To account for the interaction between elementary particles of different nature (e.g. photons and electrons), the various fields are combined into an *inseparable whole* by including coupling terms between the different fields (e.g. Maxwell's field and Dirac's field). In the usual theory, particles of the same nature (e.g. two electrons) interact only indirectly via their interaction with particles of other nature (e.g. photons) in accordance with a similar idea in the classical theory, namely, the action through a medium. In what follows, I shall always consider the whole field, including the interactions.

Experimentally, all phenomena concerning the interaction of the elementary particles are invariably studied in a volume which (being at least of atomic dimensions) is practically infinite in comparison with any effective volume of the elementary particles (as derived from the observed cross-sections of their collisions). Hence, for simplicity, one may suppose the whole field to be periodic, the basis being an *arbitrarily* large 'fundamental' cube of sides L, because this does not affect the asymptotic results for large values of L.

The fundamental property of a field is, in the classical theory, the localization of the energy, momentum, and charge. This arises from the fact that, as the consequences of the field equations, an energy-momentum-stress tensor T^{ν}_{μ} and a charge-current vector j^{ν} exist at every point $\mathbf{x} = (x^1, x^2, x^3)$ and time $t = x^0$ which obey the equations of continuity

$$\sum_{\nu=0}^{3} \frac{\partial T^{\nu}_{\mu}}{\partial x^{\nu}} = 0 \quad (\mu = 0, 1, 2, 3), \qquad \sum_{\nu=0}^{3} \frac{\partial j^{\nu}}{\partial x^{\nu}} = 0. \tag{1·1}$$

By integrating these equations over the fundamental cube, one obtains the conservation laws of the total energy H, the total momentum $\mathbf{G} = (G_1, G_2, G_3)$, and the total charge Q contained in the cube

$$H = \iiint T^4_4 dx^1 dx^2 dx^3, \quad G_1 = \iiint T^4_1 dx^1 dx^2 dx^3, \quad \ldots, \quad Q = \iiint j^4 dx^1 dx^2 dx^3, \tag{1·2}$$

which are therefore constant in time. If the field equations can be derived from a variational principle, it can be shown that T^4_1, T^4_2, T^4_3 and j^4 are always bilinear expressions in the field quantities, while T^4_4 contains besides bilinear terms also terms of higher degree (Pauli 1941, part I, § 2). This is also the case if some field equation has to be regarded as supplementary to the variational equations, e.g. in the case of Maxwell's field and Dirac's field in interaction (Born & Peng 1944, § 3).

Let H^0, H^1, ... denote the parts of H which arise from the bilinear, trilinear, ... terms of T_4^4 respectively,

$$H^0 = \iiint (\text{bilinear terms of } T_4^4)\, dx^1 dx^2 dx^3, \tag{1·3}$$

$$H^1 = \iiint (\text{trilinear terms of } T_4^4)\, dx^1 dx^2 dx^3, \dots \tag{1·4}$$

In general, any field quantity (say f) is complex; real field quantities may be considered as special cases. By a three-dimensional Fourier transformation of all the field quantities,

$$f(\mathbf{x},t) = L^{-\frac{3}{2}} \sum_k f_k(t)\, e^{i\mathbf{k}.\mathbf{x}}, \quad f^*(\mathbf{x},t) = L^{-\frac{3}{2}} \sum_k f_k^*(t)\, e^{-i\mathbf{k}.\mathbf{x}}, \tag{1·5}$$

where the summation of the wave vectors $\mathbf{k}$ extends over the lattice points of the infinite k-lattice (a cubic lattice of the lattice constant L^{-1}), the volume integration in H^0, $\mathbf{G}$ and Q is transformed into the lattice summation of the wave vectors. The contribution from each wave vector $\mathbf{k}$, being a Hermitian bilinear form of the corresponding f_k's and f_k^*'s of all the field quantities, can further be reduced to the canonical form by a linear transformation. One thus obtains

$$H^0 = \sum_k \sum_\sigma \epsilon_{k\sigma} a_{k\sigma} a_{k\sigma}^*, \quad \mathbf{G} = \sum_k \sum_\sigma \mathbf{p}_{k\sigma} a_{k\sigma} a_{k\sigma}^*, \quad Q = \sum_k \sum_\sigma e_\sigma a_{k\sigma} a_{k\sigma}^*, \tag{1·6}$$

where there are as many $a_{k\sigma}$ and $a_{k\sigma}^*$, distinguished by the index σ, as the number of linearly independent f_k's and f^*'s. The normalization factor $L^{-\frac{3}{2}}$ has been introduced in (1·5) to destroy the factor L^3 arising from the volume integration so that H^0, $\mathbf{G}$, and Q contain no factor in L. Since each additional field quantity brings in a factor $L^{-\frac{3}{2}}$, one has, for H^1, H^2, ...:

$$H^1 = L^{-\frac{3}{2}} \,(\text{trilinear form of all } a_{k\sigma} \text{ and } a_{k\sigma}^*), \tag{1·7}$$

$$H^2 = L^{-3} \,(\text{quadrilinear form of all } a_{k\sigma} \text{ and } a_{k\sigma}^*), \dots \tag{1·8}$$

where the a's and a^*'s of different $k\sigma$ are intermingled in some definite way. In the usual theory, H^1 is linear in such natural constants as the elementary charge or the mesonic charge of a nucleon, while H^2 is bilinear in these interaction constants, and so on. The series of terms of increasingly higher degree in the a's and a^*'s,

$$H = H^0 + H^1 + H^2 + \dots, \tag{1·9}$$

can thus be regarded as an expansion of H with respect to the interaction constants.

In the quantum theory, the field quantities are replaced by q-number quantities obeying non-commutative algebra, while the four co-ordinates $\mathbf{x}$, t are retained as c-number quantities. The Fourier transformation (1·5) can be used for q-number field quantities $f(\mathbf{x},t)$ and q-number Fourier amplitudes $f_k(t)$ where the wave vectors $\mathbf{k}$ are retained like $\mathbf{x}$ as c-number quantities. The $a_{k\sigma}$ and $a_{k\sigma}^*$, which are the linearly independent combinations of all the $f_k(t)$ and $f_k^*(t)$, are now q-number quantities. The total energy H of the system—namely, the whole field in the fundamental cube

—plays the role of the Hamiltonian. The quantum conditions between the a's and the a^*'s are determined by the condition that the quantum-mechanical equations of motion in the Heisenberg representation, viz.

$$\frac{\hbar}{i}\frac{\partial a_{k\sigma}}{\partial t} = Ha_{k\sigma} - a_{k\sigma}H, \quad \frac{\hbar}{i}\frac{\partial a_{k\sigma}^*}{\partial t} = Ha_{k\sigma}^* - a_{k\sigma}^*H, \tag{1·10}$$

should be formally the same as those obtained from the field equations in the classical theory. (The asterisk, when used with a q-number quantity, denotes the adjoint.) In the linear transformation from the f_k and f_k^* to the $a_{k\sigma}$ and $a_{k\sigma}^*$ it is convenient to introduce a factor $\hbar^{-\frac{1}{2}}$ so that the quantum conditions between the a's and the a^*'s no longer contain $\hbar$, viz.

$$\left.\begin{aligned} a_{k\sigma}^* a_{k'\sigma'}^* \mp a_{k'\sigma'}^* a_{k\sigma}^* = a_{k\sigma} a_{k'\sigma'} \mp a_{k'\sigma'} a_{k\sigma} = 0,\\ a_{k\sigma}^* a_{k'\sigma'} \mp a_{k'\sigma'} a_{k\sigma}^* = \delta_{kk'}\delta_{\sigma\sigma'}.\end{aligned}\right\} \tag{1·11}$$

The + sign (anti-commutation laws) is used only if *both* $a_{k\sigma}$ (or its adjoint) and $a_{k'\sigma'}$ (or its adjoint) are derived from field quantities which transform for the Lorentz transformation like spinors. The − sign (commutation laws), on the other hand, is used where at least one of these quantities is derived from field quantities which transform for the Lorentz transformation like tensors (including vectors and scalars).

As consequences of the quantum conditions, the q-number quantities $a_{k\sigma}a_{k\sigma}^*$ for various $k\sigma$ all commute. Their eigenvalues $N_{k\sigma}$ are independently $0, 1, 2, \ldots$ in the case of commutation laws being used, or 0 and 1 in the case of anti-commutation laws being used. Since G_1, G_2, G_3, Q and H^0 of (1·6) are linear combinations of the commuting quantities $a_{k\sigma}a_{k\sigma}^*$ with c-number coefficients, they all commute. In fact, $\mathbf{G}$ and Q commute with the Hamiltonian H, which expresses the conservation laws of the total momentum and charge in quantum theory. In the following, the total momentum and charge of the system will be assumed to be known.

After having obtained the quantum conditions in the Heisenberg representation, it is practical to change to the Schrödinger representation in which all q-number quantities are represented by linear operators constant in time (Dirac 1935, §§ 31–32). The operand Ψ, called the wave function, describes the quantum *state* of the system and varies with time according to Schrödinger's wave equation in time†

$$-\frac{\hbar}{i}\frac{\partial \Psi}{\partial t} = H\Psi, \quad \text{where } \Psi \text{ is short for } \Psi(N_{k\sigma}; t). \tag{1·12}$$

---

† The treatment described in the text is the usual short-cut way of applying Heisenberg and Pauli's method of quantization in the k-representation. The complete treatment, which includes also the quantum-mechanical equations of motion in space, can easily be reduced to the above simple treatment (cf. Born & Peng, 1944, §4). In the complete treatment, the wave function $\Psi(N_{k\sigma}; \mathbf{x}, t)$ satisfies the wave equations in time and space,

$$-\frac{\hbar}{i}\frac{\partial \Psi(N_{k\sigma}; \mathbf{x}, t)}{\partial t} = H\Psi(N_{k\sigma}; \mathbf{x}, t), \quad \frac{\hbar}{i}\,\mathrm{grad}\,\Psi(N_{k\sigma}; \mathbf{x}, t) = \mathbf{G}\Psi(N_{k\sigma}; \mathbf{x}, t).$$

If the total momentum is known, say $\mathbf{G}'$, the complete treatment reduces to the simple treatment by putting

$$\Psi(N_{k\sigma}; \mathbf{x}, t) = L^{-\frac{3}{2}} e^{i\mathbf{G}'\cdot\mathbf{x}/\hbar}\Psi(N_{k\sigma}; t).$$

The eigenvalues $N_{k\sigma}$ of $a_{k\sigma}a^*_{k\sigma}$ for all $k\sigma$ are adopted as the arguments of Ψ on which all $a_{k\sigma}$ and $a^*_{k\sigma}$ operate. For a set of basic functions of all $N_{k\sigma}$ one can use the simultaneous eigenfunctions of all $a_{k\sigma}a^*_{k\sigma}$, viz.

$$\phi_n^{(0)}(N_{k\sigma}) = \prod_k \prod_\sigma \delta_{n_{k\sigma}N_{k\sigma}}, \tag{1·13}$$

where n denotes in short a set of quantum numbers $\{n_{k\sigma}\}$ with all $k\sigma$. These functions, for all n, form a complete system of linearly independent functions of the N's, and are orthonormal:

$$(\phi_n^{(0)}\cdot\phi_m^{(0)}) \equiv \sum_{N_{k\sigma}} \phi_n^{(0)*}(N_{k\sigma})\,\phi_m^{(0)}(N_{k\sigma}) = \prod_k \prod_\sigma \delta_{n_{k\sigma}m_{k\sigma}} \equiv \delta_{nm}. \tag{1·14}$$

As the zero-order approximation, it is convenient to introduce an idealized system, which is described by the simpler Hamiltonian H^0, with the same $\mathbf{G}$ and Q as before, see (1·6). Evidently, the functions (1·13) form a set of eigenfunctions for this idealized system:

$$H^0\phi_n^{(0)} = E_n^{(0)}\phi_n^{(0)}, \quad \mathbf{G}\phi_n^{(0)} = \mathbf{G}_n\phi_n^{(0)}, \quad Q\phi_n^{(0)} = Q_n\phi_n^{(0)}, \tag{1·15}$$

where $\phi_n^{(0)}$ is short for $\phi_n^{(0)}(N_{k\sigma})$, with the eigenvalues

$$E_n^{(0)} = \sum_k \sum_\sigma \epsilon_{k\sigma} n_{k\sigma}, \quad \mathbf{G}_n = \sum_k \sum_\sigma \mathbf{p}_{k\sigma} n_{k\sigma}, \quad Q_n = \sum_k \sum_\sigma e_\sigma n_{k\sigma}. \tag{1·16}$$

The solution $\Phi^{(0)}$, short for $\Phi^{(0)}(N_{k\sigma};t)$, of Schrödinger's wave equation for the idealized system is simple; the general solution being a superposition of the stationary state solutions $\Phi_n^{(0)}$ with arbitrary constants:

$$-\frac{\hbar}{i}\frac{\partial\Phi^{(0)}}{\partial t} = H^0\Phi^{(0)}, \quad \Phi_n^{(0)} = \exp\left[-iE_n^{(0)}t/\hbar\right]\phi_n^{(0)}(N_{k\sigma}). \tag{1·17}$$

The peculiar form of the eigenvalues (1·16) suggests the usual interpretation that, for the state $\phi_n^{(0)}$ characterized by the set of quantum numbers $\{n_{k\sigma}\}$, there are $n_{k\sigma}$ *quanta* of the kind $k\sigma$ within the fundamental cube, each quantum of the kind $k\sigma$ possessing the energy $\epsilon_{k\sigma}$, the momentum $\mathbf{p}_{k\sigma}$, and the charge e_σ. The index σ combines the description of the nature of the quantum (electron-*like*, photon-*like*, etc.) with its spin orientations ($2s+1$ orientations for a quantum of spin s) and electric charge (positive, negative or neutral according to $e_\sigma = +e$, $-e$, or 0), while the index k describes the momentum of the quantum ($\mathbf{p}_{k\sigma} = \pm\hbar\mathbf{k}$ if $e_\sigma = \pm e$, $\mathbf{p}_{k\sigma} = \hbar\mathbf{k}$ if $e_\sigma = 0$). The energy $\epsilon_{k\sigma}$ of the quantum, of the form $+\sqrt{(c^2p^2_{k\sigma}+m^2_\sigma c^4)}$, is positive and independent of the spin orientations of the quantum and the directions of the momentum of the quantum. (It depends, however, on the nature of the quantum through the rest-energy $m_\sigma c^2$ of the quantum.) It follows from this interpretation that, excepting only the vacuum-*like* state with all $n_{k\sigma} = 0$, one can also specify the state $\phi_n^{(0)}$ merely by the numbers $n_1, n_2, \ldots, n_K$ of the quanta of various kinds $(1, 2, \ldots, K)$ which really occur (i.e. $n_{k\sigma} \neq 0$ for $k\sigma = 1, 2, \ldots, K$), if at the same time the σ's and the $\mathbf{k}$'s of these quanta are given, say $\sigma_1, \sigma_2, \ldots, \sigma_K$ and $\mathbf{k}_1, \ldots, \mathbf{k}_{K-1}$ ($\mathbf{k}_K$ can then be inferred from the total momentum).

It must be remembered that the quanta of the system H^0 represent an idealization of the real elementary particles as observed in nature. The latter are described by the Hamiltonian H, including the interaction terms H^1, H^2, etc. In these terms, $a_{k\sigma}$ and $a^*_{k\sigma}$ do not occur in the simple combination $a_{k\sigma}a^*_{k\sigma}$. The effect of $a_{k\sigma}$ or $a^*_{k\sigma}$ operating on a function of all the N's is, apart from a numerical factor depending on the N's, to increase or decrease the argument $N_{k\sigma}$ by unity (the result vanishes if $N_{k\sigma}+1$ or $N_{k\sigma}-1$ is no longer an eigenvalue of $a_{k\sigma}a^*_{k\sigma}$). The effect of $a_{k\sigma}$ or $a^*_{k\sigma}$ operating on a basic function $\phi^{(0)}_n$ is thus to produce, if the result does not vanish, a numerical multiple of a new basic function which corresponds to one quantum more or one less of the kind $k\sigma$ in the fundamental cube. From this, the effect of H^1, H^2, etc., on a function of all the N's can be deduced.

Any state of affairs concerning the elementary particles is described by a corresponding solution of the wave equation (1·12) for H (not (1·17) for the quanta). In practice, only effects of stationary processes are studied, such as the cross-sections for the collisions of the elementary particles, and also their self-energies. Hence, throughout this paper, *only stationary*-state (or quasi-stationary-state) solutions of (1·12) are considered. These are of the form

$$\Psi_a = e^{-iE_a t/\hbar}\,\psi_a(N_{k\sigma}), \tag{1·18}$$

where the ψ_a and the E_a are the eigenfunctions and the eigenvalues of the Hamiltonian H,

$$H\psi_a = \psi_a E_a, \quad H = H^0+H^1+H^2+\dots. \tag{1·19}$$

As the ψ_a are functions of all the N's, they can be expressed, when normalized, with the help of a unitary transformation of the basic functions $\{\phi^{(0)}_n\}$, say

$$\psi_a = \sum_m \phi^{(0)}_m(\phi^{(0)}_m \,.\, \psi_a). \tag{1·20}$$

Substitute this into (1·19) and develop $H\phi^{(0)}_m$ in the basis $\{\phi^{(0)}_n\}$. Then (1·19) is transformed into the system of algebraic equations

$$\sum_m (\phi^{(0)}_n \,.\, H\phi^{(0)}_m)(\phi^{(0)}_m \,.\, \psi_a) = (\phi^{(0)}_n \,.\, \psi_a)\,E_a, \tag{1·21}$$

$$(\phi^{(0)}_n \,.\, H\phi^{(0)}_m) = (\phi^{(0)}_n \,.\, H^0\phi^{(0)}_m) + (\phi^{(0)}_n \,.\, H^1\phi^{(0)}_m) + (\phi^{(0)}_n \,.\, H^2\phi^{(0)}_m) + \dots, \tag{1·22}$$

with, according to (1·15) and (1·14),

$$(\phi^{(0)}_n \,.\, H^0\phi^{(0)}_m) = \delta_{nm}E^{(0)}_m. \tag{1·23}$$

The solution of (1·21) can be successively approximated by the perturbation method, by using (1·23) as the zero-order approximation.

For physical interpretation it is sometimes convenient to analyse the stationary-state solution (1·18) of the actual system into a superposition of the stationary-state solutions (1·17) of the idealized system, say

$$\Psi_a = \sum_m \Phi^{(0)}_m(\Phi^{(0)}_m \,.\, \Psi_a), \tag{1·24}$$

with the coefficients $(\Phi_m^{(0)} . \Psi_a)$ varying with time,

$$(\Phi_m^{(0)} . \Psi_a) = \exp[i(E_m^{(0)} - E_a)\, t/\hbar]\, (\phi_m^{(0)} . \psi_a). \tag{1·25}$$

Substitute (1·24) into Schrödinger's wave equation (1·12) and use (1·17). Then (1·12) is transformed into the system of ordinary differential equations

$$\begin{aligned} -\frac{\hbar}{i}\frac{d(\Phi_n^{(0)} . \Psi_a)}{dt} &= \sum_m (\Phi_n^{(0)} . [H^1 + H^2 + \ldots]\, \Phi_m^{(0)})\, (\Phi_m^{(0)} . \Psi_a) \\ &= \sum_m \exp[i(E_n^{(0)} - E_m^{(0)})\, t/\hbar]\, \{(\phi_n^{(0)} . H^1 \phi_m^{(0)}) + (\phi_n^{(0)} . H^2 \phi_m^{(0)}) + \ldots\}\, (\Phi_m^{(0)} . \Psi_a). \end{aligned} \tag{1·26}$$

The perturbation treatment for the successive approximation to the solution of (1·26) in the *time representation* is of course mathematically connected with that of (1·21) in the *energy representation*.

2. DEGENERACY—THE FAILURE OF THE EXPANSION METHOD

In order to explain the failure of the expansion method, it will be shown that the unperturbed system H^0 is degenerate. For this purpose it is only necessary to consider the states $\phi_n^{(0)}$ possessing equal eigenvalues for the total momentum **G** and equal eigenvalues for the total charge Q, because the states possessing different eigenvalues for **G** and/or Q fall into different non-combining classes.

The vacuum-like state, which belongs only to the class of zero total momentum and zero total charge, is, of course, non-degenerate. The states which correspond to but one kind of quanta in the fundamental cube ($K = 1$) are specified by n_1 and σ_1 only. These are in general spin-degenerate (meaning that different states with different spin orientations of the quanta exist and possess precisely equal energy values), and can be partitioned into a number of mutually exclusive groups of rigorously degenerate states.

For the states which are specified by $n_1, n_2, \ldots, n_K$; $\sigma_1, \sigma_2, \ldots, \sigma_K$; and $\mathbf{k}_1, \ldots, \mathbf{k}_{K-1}$ with $K \geqslant 2$, since the k's can vary almost continuously (namely, by units of L^{-1}), the energy values form an almost continuous spectrum for each specification of the n's and σ's. In such circumstances, all the states $\phi_m^{(0)}$ with $E_m^{(0)}$ lying between

$$E_n^{(0)} - \tfrac{1}{2}\Delta E_n^{(0)} \quad \text{and} \quad E_n^{(0)} + \tfrac{1}{2}\Delta E_n^{(0)}$$

can be said to be approximately degenerate with respect to the central state $\phi_n^{(0)}$ within the allowance $\Delta E_n^{(0)}$. The allowance $\Delta E_n^{(0)}$ has to be large in comparison to the spacing of the energy values, but is still of the order of L^{-1} (i.e. $L \Delta E_n^{(0)}$ is independent of L). Thus for sufficiently large values of L, the allowance $\Delta E_n^{(0)}$ can be made arbitrarily small, and hence the approximate degeneracy becomes *almost* rigorous. The totality of such states (including the state $\phi_n^{(0)}$ itself) will be denoted by $\deg^{(0)} n$. The notation $m \in \deg^{(0)} n$ and $m \bar{\in} \deg^{(0)} n$ will be used to denote that $\phi_m^{(0)}$ is or is not almost degenerate with respect to $\phi_n^{(0)}$. Note that, if $m \in \deg^{(0)} n$, then the set $\deg^{(0)} n$ partly overlaps with the set $\deg^{(0)} m$; it is therefore impossible to partition the states of an almost continuous spectrum into a number of mutually exclusive sets of almost rigorously degenerate states.

The notation $\deg^{(0)} n$ may also be used if the state $\phi^{(0)}_n$ corresponds to but one kind of quanta in the cube ($K = 1$), then $\deg^{(0)} m$ coincides with $\deg^{(0)} n$ if $m \in \deg^{(0)} n$. For a non-degenerate state $\phi^{(0)}_n$, $\deg^{(0)} n$ consists of only the state $\phi^{(0)}_n$ itself.

The degeneracy of the states of the unperturbed system H^0 usually plays no role for the first-order perturbation, because $(\phi^{(0)}_m . H^1\phi^{(0)}_n)$ usually vanishes for $m \in \deg^{(0)} n$. (Since H^1 is trilinear in the operators $a_{k\sigma}$ and $a^*_{k\sigma}$, $(\phi^{(0)}_m . H^1\phi^{(0)}_n)$ differs from zero only if $n_{k\sigma} = m_{k\sigma} \pm 1$ simultaneously for three kinds of quanta, two of which are usually of the same nature. Then it is kinematically impossible to satisfy the conservation laws of both the momentum and the energy, except in the case where electromagnetic quanta of arbitrarily small energy and momentum are concerned. This exception arises from the fact that the rest mass of the electromagnetic quanta vanishes!) But, in any case, as will appear immediately, the degeneracy of the states of H^0 cannot be ignored from the second-order perturbation onwards.

The failure of the expansion method can now be demonstrated by applying the important theorem well known in the perturbation theory for degenerate systems (cf. Born, Heisenberg & Jordan 1926; Schrödinger 1926; Born & Jordan 1930, chap. v), which in these works is expressed in the energy representation (1·19) or (1·21).

THEOREM. If some states of the unperturbed system are degenerate, whether rigorously degenerate or almost so, then the eigenfunctions ψ_a of the perturbed system cannot be expanded straightforwardly with respect to the perturbation by a power series of the form

$$\psi_a = \phi^{(0)}_a + \psi^1_a + \psi^2_a + \dots, \tag{2·1}$$

where ψ^1_a contains the first power of the interaction constant as a factor, ψ^2_a contains the second power of the interaction constant as a factor, and so on.

Proof. This theorem can easily be proved by contradiction, as follows. If the expansion (2·1) were possible, (1·21) and (1·22) show that E_a can be similarly expanded, so (1·21) becomes, by using also (1·23),

$$\sum_m \{\delta_{nm}E^0_m + (\phi^{(0)}_n . H^1\phi^{(0)}_m) + (\phi^{(0)}_n . H^2\phi^{(0)}_m) + \dots\}\{\delta_{ma} + (\phi^{(0)}_m . \psi^1_a) + (\phi^{(0)}_m . \psi^2_a) + \dots\}$$
$$= \{\delta_{na} + (\phi^{(0)}_n . \psi^1_a) + (\phi^{(0)}_n . \psi^2_a) + \dots\}\{E^{(0)}_a + E^1_a + E^2_a + \dots\}. \tag{2·2}$$

The zero-order approximation to (2·2) is, of course, an identity. The first approximation to (2·2) then leads to

$$(\phi^{(0)}_n . H^1\phi^{(0)}_a) = \delta_{na}E^1_a + (E^{(0)}_a - E^{(0)}_n)(\phi^{(0)}_n . \psi^1_a), \tag{2·3}$$

which gives rise to the *condition of solubility*

$$(\phi^{(0)}_n . H^1\phi^{(0)}_a) = \delta_{na}E^1_a \quad \text{for} \quad n \in \deg^{(0)} a. \tag{2·4}$$

Of interest are two cases, (i) $(\phi^{(0)}_n . H^1\phi^{(0)}_a) \neq 0$ for $n \in \deg^{(0)} a$ and $n \neq a$. Then the condition of solubility (2·4) is not satisfied, and hence the theorem is proved by contradiction. This is the case where Fermi's longitudinal quanta of arbitrarily small momentum and energy are concerned. (ii) $(\phi_n^{(0)}\,.\,H^1\phi_a^{(0)}) = 0$ for $n \in \deg^{(0)} a$. This represents the general case in quantized field theories. The condition of solubility (2·4) is then satisfied by $E_a^1 = 0$, and the solution of (2·3) gives

$$(\phi_n^{(0)}\,.\,\psi_a^1) = \frac{(\phi_n^{(0)}\,.\,H^1\phi_a^{(0)})}{E_a^{(0)} - E_n^{(0)}} \quad \text{for} \quad n \bar{\in} \deg^{(0)} a. \tag{2·5}$$

But then the second approximation to (2·2) leads to a contradiction. It leads, with the help of (2·5), to

$$(\phi_n^{(0)}\,.\,H^2\phi_a^{(0)}) + \sum_{m \bar{\in} \deg^{(0)} a} \frac{(\phi_n^{(0)}\,.\,H^1\phi_m^{(0)})(\phi_m^{(0)}\,.\,H^1\phi_a^{(0)})}{E_a^{(0)} - E_m^{(0)}} = \delta_{na} E_a^2 + (E_n^{(0)} - E_n^{(0)})(\phi_a^{(0)}\,.\,\psi_a^2); \tag{2·6}$$

but the *condition of solubility* for (2·6), namely, that, for $n \in \deg^{(0)} a$, the left-hand side of (2·6) should be diagonal, is actually not satisfied. (In fact, for $n \in \deg^{(0)} a$ and $n \neq a$, the left-hand side of (2·6) is usually called the transition matrix element from the state $\phi_n^{(0)}$ to the state $\phi_a^{(0)}$.) Hence the expansion (2·1) fails.

As a consequence, (2·6) cannot be used for the calculation of E_a^2 by putting $n = a$, even if the resultant expression for the self-energy were not divergent.

The impossibility of the expansion (2·1) in the energy representation implies, by (1·18), that it is impossible to expand the Ψ_a in the form

$$\Psi_a = \Psi_a^{(0)} + \Psi_a^1 + \Psi_a^2 + \dots. \tag{2·7}$$

Hence, in the time representation, it is impossible to integrate the system (1·26) of differential equations by the usual expansion method, in which (1·26) is decomposed, by using (2·7), into

$$-\frac{\hbar}{i}\frac{d(\Phi_n^{(0)}\,.\,\Psi_a^1)}{dt} = \sum_m \exp\,[i(E_n^{(0)} - E_m^{(0)})\,t/\hbar]\,(\phi_n^{(0)}\,.\,H^1\phi_m^{(0)})\,\delta_{ma}, \tag{2·8}$$

$$-\frac{\hbar}{i}\frac{d(\Phi_n^{(0)}\,.\,\Psi_a^2)}{dt} = \sum_m \exp\,[i(E_n^{(0)} - E_m^{(0)})\,t/\hbar]\,\{(\phi_n^{(0)}\,.\,H^1\phi_m^{(0)})\,(\Phi_m^{(0)}\,.\,\Psi_a^1) + (\phi_n^{(0)}\,.\,H^2\phi_m^{(0)})\,\delta_{ma}\}, \tag{2·9}$$

……

The cross-sections thus obtained are therefore not justified even if the contributions from the terms of higher degree were not divergent.

The failure of the expansion method for *any degenerate* system is due to the faulty assumption, (2·1), that the eigenfunctions of H should equal approximately those of H^0, because the eigenfunctions for a degenerate system can be chosen arbitrarily to some extent. In such circumstances, the degeneracy of the unperturbed system has *first* to be treated by the method of secular perturbation. Such consideration is familiar wherever the perturbation method is applied, usually in the energy representation, to problems of atomic and molecular structures. In dealing with field theories and the interaction of elementary particles, the treatment of the degeneracy has been hitherto overlooked, probably because the problem is usually presented in the time representation.

3. THE METHOD OF SECULAR PERTURBATION

In accordance with the perturbation theory developed by the authors quoted above, the arbitrariness of choosing the eigenfunctions $\phi_n^{(0)}$ of the unperturbed degenerate system has to be treated by the method of secular perturbation. The following formulation of this method is designed to cover both the states which are rigorously degenerate and the states of the almost continuous spectrum which are almost so. The essential point of this method is the following. If the eigenfunctions of the perturbed system can be expanded, to the rth approximation, in a basis $\{\phi_a^{(r)}\}$, say, which are degenerate, then a suitable change of basis from $\{\phi_a^{(r)}\}$ to $\{\phi_a^{(r+1)}\}$ must be made in order that the eigenfunctions can be expanded in the new basis to the $r+1$th approximation. In atomic and molecular problems, the basis for the first or the second approximation is usually non-degenerate, in which case the degeneracy is said to be removed by the secular change of basis. This is, however, not the case in quantized field theories where the degeneracy persists in all approximations. Hence the method of secular perturbation has to be continued indefinitely, unless the perturbation method is discontinued at a certain stage in favour of other approximation methods being used.

Let $\psi_a^{(r)}$ and $E_a^{(r)}$ denote the rth approximation to the actual eigenfunctions ψ_a and eigenvalues E_a, so

$$H^{(r)}\psi_a^{(r)} = \psi_a^{(r)}E_a^{(r)}, \quad H^{(r)} = H^0+H^1+\ldots+H^r, \tag{3·1}$$

where quantities of higher degrees in the interaction constants than the rth are to be neglected. Assume that there exists a secular problem for the rth approximation, of the form

$$K^{(r)}\phi_a^{(r)} = \phi_a^{(r)}E_a^{(r)}, \quad K^{(r)} = H^0+K^{(r)1}+\ldots+K^{(r)r}, \tag{3·2}$$

which supplies a basis $\{\phi_a^{(r)}\}$ in which the $\psi_a^{(r)}$ of (3·1) can be expanded, just to the rth approximation, say

$$\psi_a^{(r)} = \phi_a^{(r)}+\psi_a^{(r)1}+\ldots+\psi_a^{(r)r}. \tag{3·3}$$

$K^{(r)}$ may be regarded as the Hamiltonian of a secular system which is nearer to the actual system than the unperturbed system H^0. The difference of $H^{(r)}$ and $K^{(r)}$ will be called the non-secular perturbation $J^{(r)}$,

$$\left.\begin{aligned} J^{(r)} &= H^{(r)}-K^{(r)} = J^{(r)1}+\ldots+J^{(r)r},\\ J^{(r)j} &= H^j-K^{(r)j} \quad (j=1,\ldots,r). \end{aligned}\right\} \tag{3·4}$$

For the rth approximation, the expansion (3·3) of $\psi_a^{(r)}$ is in fact arranged according to the order of magnitude of the non-secular perturbations $J^{(r)j}$ $(j=1,\ldots,r)$, while the secular perturbation $K^{(r)1}+\ldots+K^{(r)r}$ is not regarded as small. (Note that the superscript (r) with the brackets indicates the order of approximations $r = 0, 1, 2, \ldots$; while the superscript j without the brackets indicates the degree of the interaction constants, $j = 1,\ldots,r$. Further, the letter ϕ is used in connexion with the secular problems, while the letter ψ is used in connexion with the actual problems.)

For the first approximation, $r = 1$, it is well known, in the case where all states can be partitioned into mutually exclusive groups of rigorously degenerate states, that the secular perturbation $K^{(1)1}$ is such that its matrix elements between two degenerate states are the same as those of H^1, while its matrix elements between states which are not degenerate with respect to each other vanish. That is,

$$\left.\begin{aligned}(\phi_c^{(0)}.K^{(1)1}\phi_b^{(0)}) &= (\phi_c^{(0)}.H^1\phi_b^{(0)}) \quad \text{for} \quad c\in\deg^{(0)}b,\\ (\phi_c^{(0)}.K^{(1)1}\phi_b^{(0)}) &= 0 \quad \text{for} \quad c\,\bar{\in}\deg^{(0)}b.\end{aligned}\right\} \tag{3·5}$$

In this case, the matrix $(\phi_c^{(0)}.K^{(1)1}\phi_b^{(0)})$ consists of step matrices (which correspond to the partition of the states into mutually exclusive groups of degenerate states). The secular change of basis from $\{\phi_a^{(0)}\}$ to $\{\phi_a^{(1)}\}$ represents a unitary transformation,

$$\phi_a^{(1)} = \sum_b \phi_b^{(0)}(\phi_b^{(0)}.\phi_a^{(1)}), \tag{3·6}$$

where the transformation matrix $(\phi_b^{(0)}.\phi_a^{(1)})$ also consists of step matrices. That is,

$$\left.\begin{aligned}(\phi_b^{(0)}.\phi_a^{(1)}) &\neq 0 \quad \text{for} \quad b\in\deg^{(0)}a,\\ (\phi_b^{(0)}.\phi_a^{(1)}) &= 0 \quad \text{for} \quad b\,\bar{\in}\deg^{(0)}a.\end{aligned}\right\} \tag{3·7}$$

The secular problem (3·2) for $r = 1$, now written in the basis $\{\phi_c^{(0)}\}$, can then be partitioned accordingly,

$$\sum_{b\in\deg^{(0)}a}(\phi_c^{(0)}.K^{(1)}\phi_b^{(0)})(\phi_b^{(0)}.\phi_a^{(1)}) = (\phi_c^{(0)}.\phi_a^{(1)})\,E_a^{(1)} \quad (c\in\deg^{(0)}a) \tag{3·8}$$

with
$$(\phi_c^{(0)}.K^{(1)}\phi_b^{(0)}) = \delta_{cb}E_b^{(0)} + (\phi_c^{(0)}.K^{(1)1}\phi_b^{(0)}). \tag{3·9}$$

One thus expects that, in the case of the almost rigorous degeneracy for the states of the almost continuous spectrum, one can use (3·5)–(3·9) in the following way; namely, for the determination of $\phi_a^{(1)}$ and $E_a^{(1)}$ for a particular state, one considers the corresponding group $\deg^{(0)}a$ of states which are degenerate with respect to $\phi_a^{(0)}$ within an arbitrarily small allowance $\Delta E_a^{(0)}$. But, since it is now impossible to partition the states of the almost continuous spectrum into mutually exclusive groups of degenerate states, the matrix $(\phi_c^{(0)}.K^{(1)1}\phi_b^{(0)})$ now consists of some kind of band placed along the diagonal, outside of which the elements rapidly diminish and practically vanish. Since the number of states in the group $\deg^{(0)}a$ is large (and increases with L), the band covers a large number of diagonals on both sides of the principal diagonal (and even more so for larger values of L). The matrix $(\phi_c^{(0)}.\phi_b^{(1)})$ is of a similar structure.

The choice (3·5) of the secular perturbation $K^{(1)1}$ follows essentially from the condition that, in the new basis $\{\phi_b^{(1)}\}$, the eigenfunctions $\psi_a^{(1)}$ can be expanded to the first approximation in accordance with (3·3),

$$\psi_a^{(1)} = \sum_b \phi_b^{(1)}(\phi_b^{(1)}.\psi_a^{(1)}), \quad (\phi_b^{(1)}.\psi_a^{(1)}) = \delta_{ba} + (\phi_b^{(1)}.\psi_a^{(1)1}). \tag{3·10}$$

In order to show this let us regard $H^{(1)}$ as the Hamiltonian of the perturbed system, $K^{(1)}$ as that of the unperturbed system, and treat the non-secular perturbation

$J^{(1)} = J^{(1)1}$ by the expansion method. By (3·10), the eigenvalue problem (3·1) can be written in the form

$$(K^{(1)}+J^{(1)})\sum_b \phi_b^{(1)}(\phi_b^{(1)}\cdot\psi_a^{(1)}) = \sum_c \phi_c^{(1)}(\phi_c^{(1)}\cdot\psi_a^{(1)})\,E_a^{(1)}, \tag{3·11}$$

which, by using (3·2) and developing $J^{(1)}\phi_b^{(1)}$ in the basis $\{\phi_c^{(1)}\}$, becomes

$$\sum_b (\phi_c^{(1)}\cdot J^{(1)}\phi_b^{(1)})(\phi_b^{(1)}\cdot\psi_a^{(1)}) = (E_a^{(1)}-E_c^{(1)})(\phi_c^{(1)}\cdot\psi_a^{(1)}). \tag{3·12}$$

On the right-hand side of (3·12), since the leading term δ_{ca} of $(\phi_c^{(1)}\cdot\psi_a^{(1)})$ contributes nothing, the $(E_a^{(1)}-E_c^{(1)})$ may be approximated by $(E_a^{(0)}-E_c^{(0)})$. By using the expansions (3·3) and (3·4), (3·12) gives

$$(\phi_c^{(1)}\cdot J^{(1)1}\phi_a^{(1)}) = (E_a^{(0)}-E_c^{(0)})(\phi_c^{(1)}\cdot\psi_a^{(1)1}). \tag{3·13}$$

The condition of solubility for (3·13) demands

$$\left.\begin{aligned} &(\phi_c^{(1)}\cdot J^{(1)1}\phi_a^{(1)}) = 0 \quad \text{for} \quad c\in\deg^{(0)}a,\\ \text{i.e.}\qquad &(\phi_c^{(1)}\cdot K^{(1)1}\phi_a^{(1)}) = (\phi_c^{(1)}\cdot H^1\phi_a^{(1)}) \quad \text{for} \quad c\in\deg^{(0)}a.\end{aligned}\right\} \tag{3·14}$$

When expressed in the basis $\{\phi_a^{(0)}\}$, this gives the first part of (3·5). The second part of (3·5) has been chosen for simplicity and definiteness, but it is in principle arbitrary.

The majority of $E_a^{(1)}$, as determined by solving (3·8), will form an almost continuous spectrum. The majority of $\phi_a^{(1)}$ are therefore degenerate. For the second approximation, $r = 2$, a further secular change of basis from $\{\phi_a^{(1)}\}$ to $\{\phi_a^{(2)}\}$ is therefore necessary, say

$$\phi_a^{(2)} = \sum_b \phi_b^{(1)}(\phi_b^{(1)}\cdot\phi_a^{(2)}), \tag{3·15}$$

where

$$\left.\begin{aligned} (\phi_b^{(1)}\cdot\phi_a^{(2)}) &\neq 0 \quad \text{for} \quad b\in\deg^{(1)}a,\\ (\phi_b^{(1)}\cdot\phi_a^{(2)}) &= 0 \quad \text{for} \quad b\notin\deg^{(1)}a.\end{aligned}\right\} \tag{3·16}$$

Here $\deg^{(1)}a$ denotes a group of states $\phi_c^{(1)}$ with $E_c^{(1)}$ lying within an allowance $\Delta E_a^{(1)}$ around the central value $E_a^{(1)}$. The secular perturbation for the determination of the new basis $\{\phi_a^{(2)}\}$ can be derived similarly by treating the non-secular perturbation $J^{(2)}$ for the second approximation by the expansion method, as follows. In the basis $\{\phi_c^{(2)}\}$, (3·1) becomes, by (3·2),

$$\sum_b (\phi_c^{(2)}\cdot J^{(2)}\phi_b^{(2)})(\phi_b^{(2)}\cdot\psi_a^{(2)}) = (E_a^{(2)}-E_c^{(2)})(\phi_c^{(2)}\cdot\psi_a^{(2)}). \tag{3·17}$$

By using the expansions (3·3) and (3·4), and approximating the $(E_a^{(2)}-E_c^{(2)})$ by $(E_a^{(1)}-E_c^{(1)})$, (3·17) gives, for the first-order and the second-order quantities,

$$(\phi_c^{(2)}\cdot J^{(2)1}\phi_a^{(2)}) = (E_a^{(1)}-E_c^{(1)})(\phi_c^{(2)}\cdot\psi_a^{(2)1}), \tag{3·18}$$

$$(\phi_c^{(2)}\cdot J^{(2)2}\phi_a^{(2)})+\sum_b (\phi_c^{(2)}\cdot J^{(2)1}\phi_b^{(2)})(\phi_b^{(2)}\cdot\psi_a^{(2)1}) = (E_a^{(1)}-E_c^{(1)})(\phi_c^{(2)}\cdot\psi_a^{(2)2}). \tag{3·19}$$

The condition of solubility of (3·18) determines $K^{(2)1}$, the first-order secular perturbation in the second approximation, viz.

$$\left.\begin{aligned} &(\phi_c^{(2)}\,.\,J^{(2)1}\phi_a^{(2)}) = 0 \quad \text{for} \quad c \in \deg^{(1)} a,\\ \text{i.e.}\qquad &(\phi_c^{(2)}\,.\,K^{(2)1}\phi_a^{(2)}) = (\phi_c^{(2)}\,.\,H^1\phi_a^{(2)}) \quad \text{for} \quad c \in \deg^{(1)} a. \end{aligned}\right\} \tag{3·20}$$

For simplicity and definiteness, take

$$\left.\begin{aligned} &(\phi_c^{(2)}\,.\,K^{(2)1}\phi_a^{(2)}) = 0 \quad \text{for} \quad c \,\bar{\in}\, \deg^{(1)} a,\\ \text{so}\qquad &(\phi_c^{(2)}\,.\,J^{(2)1}\phi_a^{(2)}) = (\phi_c^{(2)}\,.\,H^1\phi_a^{(2)}) \quad \text{for} \quad c \,\bar{\in}\, \deg^{(1)} a. \end{aligned}\right\} \tag{3·21}$$

The solution of (3·18) then gives

$$(\phi_c^{(2)}\,.\,\psi_a^{(2)1}) = \frac{(\phi_c^{(2)}\,.\,H^1\phi_a^{(2)})}{E_a^{(1)} - E_c^{(1)}} \quad \text{for} \quad c \,\bar{\in}\, \deg^{(1)} a, \tag{3·22}$$

and, for simplicity and definiteness, take

$$(\phi_c^{(2)}\,.\,\psi_a^{(2)1}) = 0 \quad \text{for} \quad c \in \deg^{(1)} a. \tag{3·23}$$

With the help of these relations, the condition of solubility for (3·19), namely,

$$(\phi_c^{(2)}\,.\,J^{(2)2}\phi_a^{(2)}) + \sum_b (\phi_c^{(2)}\,.\,J^{(2)1}\phi_b^{(2)})\,(\phi_b^{(2)}\,.\,\psi_a^{(2)1}) = 0 \quad \text{for} \quad c \in \deg^{(1)} a, \tag{3·24}$$

determines the second-order secular perturbation $K^{(2)2}$,

$$(\phi_c^{(2)}\,.\,K^{(2)2}\phi_a^{(2)}) = (\phi_c^{(2)}\,.\,H^2\phi_a^{(2)}) + \sum_{b \,\bar{\in}\, \deg^{(1)} a} \frac{(\phi_c^{(2)}\,.\,H^1\phi_b^{(2)})\,(\phi_b^{(2)}\,.\,H^1\phi_a^{(2)})}{E_a^{(1)} - E_b^{(1)}} \quad \text{for} \quad c \in \deg^{(1)} a. \tag{3·25}$$

For definiteness, take

$$\left.\begin{aligned} &(\phi_c^{(2)}\,.\,K^{(2)2}\phi_a^{(2)}) = 0 \quad \text{for} \quad c \,\bar{\in}\, \deg^{(1)} a,\\ \text{so}\qquad &(\phi_c^{(2)}\,.\,J^{(2)2}\phi_a^{(2)}) = (\phi_c^{(2)}\,.\,H^2\phi_a^{(2)}) \quad \text{for} \quad c \,\bar{\in}\, \deg^{(1)} a. \end{aligned}\right\} \tag{3·26}$$

The solution of (3·19) then gives

$$\begin{aligned} (\phi_c^{(2)}\,.\,\psi_a^{(2)2}) = \frac{1}{E_a^{(1)} - E_c^{(1)}} \Bigg\{ &(\phi_c^{(2)}\,.\,H^2\phi_a^{(2)}) \\ &+ \sum_{b \,\bar{\in}\, \deg^{(1)} a} \frac{(\phi_c^{(2)}\,.\,H^1\phi_b^{(2)})\,(\phi_b^{(2)}\,.\,H^1\phi_a^{(2)})}{E_a^{(1)} - E_b^{(1)}} \Bigg\} \quad \text{for} \quad c \,\bar{\in}\, \deg^{(1)} a. \end{aligned} \tag{3·27}$$

The normalization condition

$$\sum_b (\psi_c^{(2)}\,.\,\phi_b^{(2)})\,(\phi_b^{(2)}\,.\,\psi_a^{(2)}) = (\psi_c^{(2)}\,.\,\psi_a^{(2)}) = \delta_{ca} \tag{3·28}$$

leads to

$$(\psi_c^{(2)2}\,.\,\phi_a^{(2)}) + (\phi_c^{(2)}\,.\,\psi_a^{(2)2}) + \sum_b (\psi_c^{(2)1}\,.\,\phi_b^{(2)})\,(\phi_b^{(2)}\,.\,\psi_a^{(2)1}) = 0, \tag{3·29}$$

which restricts the choice of $(\phi_c^{(2)}\,.\,\psi_a^{(2)2})$ for $c \in \deg^{(1)} a$.

From (3·15) and (3·16) follow, for any continuous function $f(E_b^{(1)})$ and sufficiently small allowance $\Delta E_a^{(1)}$ used in the definition of $\deg^{(1)} a$,

$$\sum_{b \in \deg^{(1)} a} f(E_b^{(1)})\,\phi_b^{(1)}(\phi_b^{(1)}\,.\,\phi_a^{(2)}) = f(E_a^{(1)})\,\phi_a^{(2)}, \tag{3·30}$$

and conversely, as the functions $\{\phi^{(2)}\}$ are orthonormal,

$$\sum_{a\in \deg^{(1)} b} f(E_a^{(1)})\,\phi_a^{(2)}(\phi_a^{(2)}\cdot\phi_b^{(1)}) = f(E_b^{(1)})\,\phi_b^{(1)}. \tag{3·31}$$

Hence the secular problem for the determination of $\phi_a^{(2)}$ and $E_a^{(2)}$ can be written in the basis $\{\phi_c^{(1)}\}$ in the form

$$\sum_{b\in \deg^{(1)} a} (\phi_c^{(1)}\cdot K^{(2)}\phi_b^{(1)})\,(\phi_b^{(1)}\cdot\phi_a^{(2)}) = (\phi_c^{(1)}\cdot\phi_a^{(2)})\,E_a^{(2)} \quad (c\in \deg^{(1)} a), \tag{3·32}$$

where, by using the secular problem for $\phi_b^{(1)}$ and $E_b^{(1)}$,

$$\begin{aligned}(\phi_c^{(1)}\cdot K^{(2)}\phi_b^{(1)}) &= (\phi_c^{(1)}\cdot K^{(1)}\phi_b^{(1)}) + (\phi_c^{(1)}\cdot[K^{(2)}-K^{(1)}]\,\phi_b^{(1)})\\ &= \delta_{cb}\,E_b^{(1)} + (\phi_c^{(1)}\cdot[K^{(2)1}-K^{(1)1}]\,\phi_b^{(1)}) + (\phi_c^{(1)}\cdot K^{(2)2}\phi_b^{(1)}).\end{aligned} \tag{3·33}$$

By transforming (3·20) into the basis $\{\phi_b^{(1)}\}$ and comparing the result with (3·14), one sees that the correction of the first-order secular perturbation arises only from the difference of $\deg^{(1)} b$ and $\deg^{(0)} b$:

$$(\phi_c^{(1)}\cdot[K^{(2)1}-K^{(1)1}]\,\phi_b^{(1)}) = \left\{\begin{array}{lllll} (\phi_c^{(1)}\cdot H^1\phi_b^{(1)}) & \text{if} & c\in \deg^{(1)} b & \text{but} & c\,\bar{\in}\,\deg^{(0)} b,\\ -(\phi_c^{(1)}\cdot H^1\phi_b^{(1)}) & \text{if} & c\in \deg^{(0)} b & \text{but} & c\,\bar{\in}\,\deg^{(1)} b,\\ 0 \quad \text{if otherwise.} \end{array}\right\} \tag{3·34}$$

By a change of basis, (3·25) and (3·26) become

$$(\phi_c^{(1)}\cdot K^{(2)2}\phi_b^{(1)}) = (\phi_c^{(1)}\cdot H^2\phi_b^{(1)})$$
$$+\sum_{a'\,\bar{\in}\,\deg^{(1)} b} \frac{\sum_{b'\in \deg^{(1)} a'}(\phi_c^{(1)}\cdot H^1\phi_{b'}^{(1)})\,(\phi_{b'}^{(1)}\cdot\phi_{a'}^{(2)})\sum_{c'\in \deg^{(1)} a'}(\phi_{a'}^{(2)}\cdot\phi_{c'}^{(1)})\,(\phi_{c'}^{(1)}\cdot H^1\phi_b^{(1)})}{E_b^{(1)}-E_{a'}^{(1)}} \quad \text{for} \quad c\in \deg^{(1)} b, \tag{3·35}$$

and

$$(\phi_c^{(1)}\cdot K^{(2)2}\phi_b^{(1)}) = 0 \quad \text{for} \quad c\,\bar{\in}\,\deg^{(1)} b. \tag{3·36}$$

In the numerator of the fraction term of (3·35), note the occurrence of the factors arising from the secular change of basis. If the states which are included in the summation over a' can be partitioned into mutually exclusive groups of degenerate states, these factors can be eliminated by using (3·31) and the orthogonality condition of the basis $\{\phi_{b'}^{(1)}\}$. On the other hand, if the states to be summed form an almost continuous spectrum and thus cannot be partitioned into mutually exclusive groups of degenerate states, these factors remain and, as will be demonstrated later in § 6, may help to render the summation convergent.

The derivation of the secular perturbation for the higher approximations is similar. In quantized field theories, the self-energies arise only at the second, the fourth, ..., approximations. Then every two successive approximations, such as the first and the second, the third and the fourth, ..., can be dealt with in one step.

In the time representation, the method of secular perturbation assumes the following form. Let

$$\left.\begin{array}{l}\Phi_a^{(r)} = \exp[-iE_a^{(r)}t/\hbar]\,\phi_a^{(r)}, \quad \Psi_a^{(r)} = \exp[-iE_a^{(r)}t/\hbar]\,\psi_a^{(r)},\\ \Psi_a^{(r)j} = \exp[-iE_a^{(r)}t/\hbar]\,\psi_a^{(r)j} \quad (j=1,\dots,r) \quad (r=1,2,\dots).\end{array}\right\} \tag{3·37}$$

For successive approximations, one has to solve (1·26) by the successive change of variables $(r = 1, 2, \ldots)$

$$(\Phi_n^{(0)} \,.\, \Psi_a) = \sum_m \ldots \sum_b (\Phi_n^{(0)} \,.\, \Phi_m^{(1)}) \ldots (\Phi_c^{(r-1)} \,.\, \Phi_b^{(r)})\,(\Phi_b^{(r)} \,.\, \Psi_a), \tag{3·38}$$

with

$$(\Phi_c^{(r-1)} \,.\, \Phi_b^{(r)}) = \exp\,[i(E_b^{(r)} - E_c^{(r-1)})\,t/\hbar]\,(\phi_c^{(r-1)} \,.\, \phi_b^{(r)}). \tag{3·39}$$

The condition that the rth approximation to Ψ_a can be expanded in the form

$$\Psi_a^{(r)} = \Phi_a^{(r)} + \Psi_a^{(r)1} + \ldots + \Psi_a^{(r)r} \tag{3·40}$$

determines the secular problem

$$-\frac{\hbar}{i}\frac{\partial \Phi_a^{(r)}}{\partial t} = K^{(r)}\Phi_a^{(r)} \quad (r = 1, 2, \ldots). \tag{3·41}$$

The explicit expression for $K^{(r)}$ can also be derived by applying the expansion method for the treatment of the non-secular perturbation in the time representation.

From (3·41) follows the differential equations for the secular change of basis for the successive approximations, $r = 1, 2, \ldots$

$$-\frac{\hbar}{i}\frac{d(\Phi_c^{(r-1)} \,.\, \Phi_a^{(r)})}{dt} = \sum_b (\Phi_c^{(r-1)} \,.\, [K^{(r)} - K^{(r-1)}]\,\Phi_b^{(r-1)})\,(\Phi_b^{(r-1)} \,.\, \Phi_a^{(r)}). \tag{3·42}$$

For $r = 1$ or 2, (3·42) is equivalent to (3·8) or (3·32), as is easily verified by using (3·39).

4. ASYMPTOTICAL SOLUTION OF THE SECULAR PROBLEM FOR LARGE VALUES OF L—REGULAR CASE

The first-order secular perturbation generally vanishes in quantized field theories. The only exceptional case occurs where an electromagnetic quantum of arbitrarily small energy ϵ (of the order of $\hbar c L^{-1}$) is emitted or absorbed by (say) an electron-like quantum. As the corresponding matrix element varies as $\epsilon^{-\frac{1}{2}}$, this physically exceptional case represents also a mathematically singular case. Since the usual divergence difficulty originates from the quanta of arbitrarily large energies and is independent of the nature of the quanta, it is convenient, for the present general investigation, to ignore altogether the existence of the electromagnetic quanta of arbitrarily small energies, and thus to avoid any special complication peculiar to this singular case alone. The method of solution of the secular problem in the regular case will be demonstrated in detail by considering the second approximation. This method has immediate physical applications, for example, in dealing with the interaction between the nucleon field and the meson field.

Suppose that the first-order secular perturbation vanishes, so

$$\phi_a^{(1)} = \phi_a^{(0)}, \quad E_a^{(1)} = E_a^{(0)}, \quad \deg^{(1)} a = \deg^{(0)} a. \tag{4·1}$$

The second-order secular problem (3·32) becomes

$$\sum_{b \in \deg^{(0)} a} (\phi_c^{(0)} \,.\, K^{(2)}\phi_b^{(0)})\,(\phi_b^{(0)} \,.\, \phi_a^{(2)}) = (\phi_c^{(0)} \,.\, \phi_a^{(2)})\,E_a^{(2)} \quad (c \in \deg^{(0)} a), \tag{4·2}$$

where, by (3·33) and (3·35),

$$(\phi_c^{(0)}.K^{(2)}\phi_c^{(0)}) = E_c^{(0)}+(\phi_c^{(0)}.K^{(2)2}\phi_c^{(0)}) = E_c^{(0)}+(\phi_c^{(0)}.H^2\phi_c^{(0)})$$
$$+\sum_{a'\bar{\in}\deg^{(0)}c}\frac{\sum\limits_{b'\in\deg^{(0)}a'}(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.\phi_{a'}^{(2)})\sum\limits_{c'\in\deg^{(0)}a'}(\phi_{a'}^{(2)}.\phi_{c'}^{(0)})(\phi_{c'}^{(0)}.H^1\phi_c^{(0)})}{E_b^{(0)}-E_{a'}^{(0)}}, \quad (4\cdot3)$$

$$(\phi_c^{(0)}.K^{(2)}\phi_b^{(0)}) = (\phi_c^{(0)}.H^2\phi_b^{(0)})$$
$$+\sum_{a'\bar{\in}\deg^{(0)}b}\frac{\sum\limits_{b'\in\deg^{(0)}a'}(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.\phi_{a'}^{(2)})\sum\limits_{c'\in\deg^{(0)}a'}(\phi_{a'}^{(2)}.\phi_{c'}^{(0)})(\phi_{c'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{a'}^{(0)}},$$
$$(c\in\deg^{(0)}b,\ c\neq b). \quad (4\cdot4)$$

The diagonal elements and the non-diagonal elements are here written out separately because the respective fraction terms are of different orders in L. In fact, the non-diagonal elements (4·4) can be approximated by the familiar expression (known in the expansion method as the matrix element for the transition)

$$(\phi_c^{(0)}.K^{(2)}\phi_b^{(0)}) = (\phi_b^{(0)}.H^2\phi_c^{(0)})+\sum_{b'}\frac{(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{b'}^{(0)}} \quad (c\in\deg^{(0)}b,\ c\neq b), \quad (4\cdot5)$$

where the summation automatically covers only the group of intermediate states (for which the summand differs from zero). This approximation follows from the fact that, in the circumstances, the intermediate states are finite in number, possess different energy values apart from spin degeneracy, and hence can be regarded as the central states of a few mutually exclusive groups of degenerate states, say $\deg^{(0)}1$, $\deg^{(0)}2$, ..., $\deg^{(0)}f$. By using (3·31), the main contribution to the fraction terms of (4·4) is easily seen to be

$$\left\{\sum_{a'\in\deg^{(0)}1}+\sum_{a'\in\deg^{(0)}2}+\ldots+\sum_{a'\in\deg^{(0)}f}\right\}\frac{\sum\limits_{b'\in\deg^{(0)}a'}(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.\phi_{a'}^{(2)})\sum\limits_{c'\in\deg^{(0)}a'}(\phi_{a'}^{(2)}.\phi_{c'}^{(0)})(\phi_{c'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{a'}^{(0)}}$$
$$=\left\{\sum_{b'\in\deg^{(0)}1}+\sum_{b'\in\deg^{(0)}2}+\ldots+\sum_{b'\in\deg^{(0)}f}\right\}\frac{(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{b'}^{(0)}}$$
$$=\sum_{b'}\frac{(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{b'}^{(0)}}, \quad (4\cdot6)$$

because for each of these mutually exclusive groups, say $\deg^{(0)}i$ $(i = 1, 2, \ldots, f)$, one has, by a change of the order of summation, and by using (3·31) together with the fact that the functions $\{\phi_{b'}^{(1)}\}$ are orthogonal,

$$\sum_{a'\in\deg^{(0)}i}\frac{\sum\limits_{b'\in\deg^{(0)}a'}(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.\phi_{a'}^{(2)})\sum\limits_{c'\in\deg^{(0)}a'}(\phi_{a'}^{(2)}.\phi_{c'}^{(0)})(\phi_{c'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{a'}^{(0)}}$$
$$=\sum_{b'\in\deg^{(0)}i}\ \sum_{c'\in\deg^{(0)}i}\frac{(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})\,\delta_{b'c'}(\phi_{c'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{b'}^{(0)}}$$
$$=\sum_{b'\in\deg^{(0)}i}\frac{(\phi_c^{(0)}.H^1\phi_{b'}^{(0)})(\phi_{b'}^{(0)}.H^1\phi_b^{(0)})}{E_b^{(0)}-E_{b'}^{(0)}}. \quad (4\cdot7)$$

The same consideration shows that the contribution due to any other group of degenerate states can be neglected in comparison to (4·6). It is seen from (1·7) and (1·8) that the non-diagonal elements (4·5) are of the order of L^{-3}. On the other hand, all states $\phi_b^{(0)}$ for which $(\phi_b^{(0)}.H^1\phi_c^{(0)}) \neq 0$ contribute to the fraction terms of (4·3). There are a few such states (arising from $\phi_c^{(0)}$ by the virtual absorption of some quanta) which can be treated as above, with contributions of the order of L^{-3}. But the main contribution arises from the large number of states (arising from $\phi_c^{(0)}$ by the virtual emission of some quanta) which cannot be partitioned into mutually exclusive groups of degenerate states, as their energy values form an almost continuous spectrum. The number of such states varies as L^3, so their total contribution to (4·3) is of the order of L^0 (i.e. independent of L); but the actual evaluation of this contribution has to be postponed until the secular change of basis for these intermediate states of the almost continuous spectrum has been found (cf. (4·28) below).

For large values of L, let a state, say $\phi_b^{(0)}$, be denoted also by $\phi_B^{(0)}(E_b^{(0)})$. Here the energy value is indicated separately, while the capital letter used as a suffix symbolizes all the remaining detailed information which is needed for the specification of the state. The specification of the number and the nature of the quanta and their spin orientations are described by discrete variables, while the specification of the directions of the momenta of these quanta and the proportion in which the total energy is shared by these quanta are described by almost continuous variables. For the solution of (4·2) it is sufficient to restrict the group $\deg^{(0)} a$ to contain only such states $\phi_b^{(0)}$ for which $(\phi_b^{(0)}.K^{(2)}\phi_a^{(0)}) \neq 0$. Let $\rho_B^{(0)}(E_b^{(0)})\,dE_b^{(0)}$ denote the number of such states with their energy values lying in the interval $dE_b^{(0)}$ around the value $E_b^{(0)}$ of the almost continuous spectrum which further conform to the detailed specification B. Then for any continuous function $f(\phi_b^{(0)})$ of $\phi_b^{(0)}$, one has *asymptotically*

$$\sum_{b\in \deg^{(0)} a} f(\phi_b^{(0)}) = \sum_B \int_{E_a^{(0)}-\frac{1}{2}\Delta E_a^{(0)}}^{E_a^{(0)}+\frac{1}{2}\Delta E_a^{(0)}} f(\phi_B^{(0)}(E_b^{(0)}))\,\rho_B^{(0)}(E_b^{(0)})\,dE_b^{(0)}. \qquad (4\cdot8)$$

$\rho_B^{(0)}(E_b^{(0)})$ contains a factor L^3. It contains also the differentials of the continuous variables describing the detailed specification B, and the summation over B implies the integration with respect to these continuous variables. As in (4·8), the equality sign will often be used for asymptotic relations.

THEOREM. The asymptotic solution of (4·2) for large values of L is of the form

$$(\phi_a^{(0)}.\phi_a^{(2)}) = 1, \quad (\phi_b^{(0)}.\phi_a^{(2)}) = 0 \quad (E_b^{(0)} = E_a^{(0)}), \qquad (4\cdot9)$$

$$(\phi_b^{(0)}.\phi_a^{(2)}) = P_L(E_a^{(0)} - E_b^{(0)})\,(\phi_b^{(0)}.S^{(2)}\phi_a^{(0)}) \quad (b \in \deg^{(0)} a,\ E_b^{(0)} \neq E_a^{(0)}). \qquad (4\cdot10)$$

Here $P_L(E_a^{(0)} - E_b^{(0)})$ denotes a continuous function of $E_b^{(0)}$ for arbitrarily large (yet finite) values of L. Among its chief properties are the following:

$$(E_a^{(0)} - E_b^{(0)})\,P_L(E_a^{(0)} - E_b^{(0)}) = 1 \quad (E_b^{(0)} \neq E_a^{(0)}), \qquad (4\cdot11)$$

$$\int_{E_a^{(0)}-\frac{1}{2}\Delta E_a^{(0)}}^{E_a^{(0)}+\frac{1}{2}\Delta E_a^{(0)}} P_L(E_a^{(0)} - E_b^{(0)})\,f(E_b^{(0)})\,dE_b^{(0)} = -i\pi f(E_a^{(0)}), \qquad (4\cdot12)$$

for any function $f(E_b^{(0)})$ which is *smooth* in the neighbourhood of $E_a^{(0)}$, namely,

$$f(E_b^{(0)}) = f(E_a^{(0)})\{1 \pm \text{quantities of the order of } L^{-1}\}$$

for $E_b^{(0)} = E_a^{(0)} \pm$ quantities of the order of L^{-1}. The other factor of (4·10), namely,

$$(\phi_b^{(0)} . S^{(2)}\phi_a^{(0)}) = (\phi_B^{(0)}(E_b^{(0)}) . S^{(2)}\phi_A^{(0)}(E_a^{(0)})) \quad (B \neq A),$$

is smooth in the neighbourhood of $E_b^{(0)} = E_a^{(0)}$, and its value at $E_b^{(0)} = E_a^{(0)}$ is to be determined from the following system of integral equations (the integration over the continuous variables describing the detailed specification B is hidden under the summation sign):

$$\left.\begin{aligned}&(\phi_C^{(0)}(E_a^{(0)}) . S^{(2)}\phi_A^{(0)}(E_a^{(0)})) = (\phi_C^{(0)}(E_a^{(0)}) . K^{(2)}\phi_A^{(0)}(E_a^{(0)}))\\ &\quad - i\pi \sum_{B\neq C,A} (\phi_C^{(0)}(E_a^{(0)}) . K^{(2)}\phi_B^{(0)}(E_a^{(0)}))\, \rho_B^{(0)}(E_a^{(0)})\, (\phi_B^{(0)}(E_a^{(0)}) . S^{(2)}\phi_A^{(0)}(E_a^{(0)})) \quad (C \neq A).\end{aligned}\right\} \tag{4·13}$$

Since $\rho_B^{(0)}(E_a^{(0)})$ contains the factor L^3, and

$$(\phi_C^{(0)}(E_a^{(0)}) . K^{(2)}\phi_B^{(0)}(E_a^{(0)})) \quad (C \neq B)$$

contains the factor L^{-3}, (4·13) shows that

$$(\phi_B^{(0)}(E_a^{(0)}) . S^{(2)}\phi_A^{(0)}(E_a^{(0)})) \quad (B \neq A)$$

contains the factor L^{-3}.

In order to verify the solution (4·9) and (4·10), substitute these into (4·2). For $c = a$, (4·2) then becomes

$$E_a^{(2)} = (\phi_a^{(0)} . K^{(2)}\phi_a^{(0)}) + \sum_{b\in \deg^{(0)} a}^{E_b^{(0)}\neq E_a^{(0)}} (\phi_a^{(0)} . K^{(2)}\phi_b^{(0)})\, P_L(E_a^{(0)} - E_b^{(0)})\, (\phi_b^{(0)} . S^{(2)}\phi_a^{(0)}). \tag{4·14}$$

By applying (4·8) and then (4·12), this gives for $E_a^{(2)}$,

$$E_a^{(2)} = (\phi_a^{(0)} . K^{(2)}\phi_a^{(0)}) - i\pi \sum_{B\neq A} (\phi_A^{(0)}(E_a^{(0)}) . K^{(2)}\phi_B^{(0)}(E_a^{(0)}))\, \rho_B^{(0)}(E_a^{(0)})\, (\phi_B^{(0)}(E_a^{(0)}) . S^{(2)}\phi_A^{(0)}(E_a^{(0)})). \tag{4·15}$$

The $i\pi$ term is of the order of L^{-3} and can be neglected in comparison with $(\phi_a^{(0)} . K^{(2)}\phi_a^{(0)})$. Further, for $c \neq a$, (4·2) can be written in the form

$$\begin{aligned}&(\phi_c^{(0)} . K^{(2)}\phi_a^{(0)})(\phi_a^{(0)} . \phi_a^{(2)}) + \sum_{b\in \deg^{(0)} a}^{b\neq c,a} (\phi_c^{(0)} . K^{(2)}\phi_b^{(0)})(\phi_b^{(0)} . \phi_a^{(2)})\\ &\qquad = (\phi_c^{(0)} . \phi_a^{(2)})\{E_a^{(2)} - (\phi_c^{(0)} . K^{(2)}\phi_c^{(0)})\} \quad (c \in \deg^{(0)} a,\ c \neq a).\end{aligned} \tag{4·16}$$

Consider here only† such states c for which $E_c^{(0)} \neq E_a^{(0)}$. Then in view of the fact that $(\phi_c^{(0)} . \phi_a^{(2)})$ contains $(\phi_c^{(0)} . S^{(2)}\phi_a^{(0)})$ which is already of the order of L^{-3}, the coefficient on the right-hand side of (4·16) can be replaced by $(E_a^{(2)} - E_c^{(2)})$, on account of (4·15). To be consistent with the approximation made above (following (3·17)) in the derivation of the matrix elements for the secular perturbation, however, this coefficient is to be approximated by $(E^{(1)}_a - E^{(1)}_c)$ (that is, by (4·1), $(E^{(0)}_a - E^{(0)}_c)$). Thus (4·16) becomes, by substitution from (4·9) and (4·10),

---

† This is sufficient: because, for large values of L, the number of states c with $E_c^{(0)} = E_a^{(0)}$ exactly is incomparably few among the enormous number of states $c \in \deg^{(0)} a$. Further, that $E_c^{(0)} = E_a^{(0)}$ do occur is due to the fact that H^0 admits the symmetry operations (rotations and reflexions) of the space; and since this symmetry is also possessed by H, such degeneracy may be left aside in the perturbation calculations.

$$(\phi^{(0)}_c\,.\,K^{(2)}\phi^{(0)}_a) + \sum^{E^{(0)}_b \neq E^{(0)}_a}_{\substack{b\neq c\\ b\,\epsilon\,\deg^{(0)} a}} (\phi^{(0)}_c\,.\,K^{(2)}\phi^{(0)}_b)\, P_L(E^{(0)}_a - E^{(0)}_b)\,(\phi^{(0)}_b\,.\,S^{(2)}\phi^{(0)}_a)$$
$$= (E^{(0)}_a - E^{(0)}_c)\, P_L(E^{(0)}_a - E^{(0)}_c)\,(\phi^{(0)}_c\,.\,S^{(2)}\phi^{(0)}_a), \quad (c\,\epsilon\,\deg^{(0)} a,\ E^{(0)}_c \neq E^{(0)}_a). \quad (4{\cdot}17)$$

By using (4·8), (4·11) and (4·12), and making the allowance $\Delta E^{(0)}_a$ arbitrarily small, (4·17) becomes (4·13). Thus the asymptotic solution (4·9) and (4·10) is verified for any state $\phi^{(2)}_a$ which is derived from the corresponding group $\deg^{(0)} a$ of the almost continuous spectrum.

The complex conjugate of (4·9) and (4·10) are

$$(\phi^{(2)}_a\,.\,\phi^{(0)}_a) = 1, \quad (\phi^{(2)}_a\,.\,\phi^{(0)}_b) = 0 \quad (E^{(0)}_b = E^{(0)}_a); \quad (4{\cdot}18)$$
$$(\phi^{(2)}_a\,.\,\phi^{(0)}_b) = P^*_L(E^{(0)}_a - E^{(0)}_b)\,(S^{(2)}\phi^{(0)}_a\,.\,\phi^{(0)}_b) \quad (b\,\epsilon\,\deg^{(0)} a,\ E^{(0)}_b \neq E^{(0)}_a). \quad (4{\cdot}19)$$

One has also to verify that the $\phi^{(2)}_a$ are orthonormal. For two different states, consider

$$(\phi^{(2)}_c\,.\,\phi^{(2)}_a) = ({}^{(2)}_{c}\phi\,.\,\phi^{(0)}_a) + (\phi^{(0)}_c\,.\,\phi^{(2)}_a) + \sum_{b\neq c,\,a} (\phi^{(2)}_c\,.\,\phi^{(0)}_b)\,(\phi^{(0)}_b\,.\,\phi^{(2)}_a) \quad (c \neq a). \quad (4{\cdot}20)$$

Of course, (4·20) vanishes if $\deg^{(0)} a$ and $\deg^{(0)} c$ do not overlap; it will be shown that (4·20) also vanishes if $c\,\epsilon\,\deg^{(0)} a$, so the two groups $\deg^{(0)} a$ and $\deg^{(0)} c$ do overlap. Suppose $E^{(0)}_c \neq E^{(0)}_a$. By substituting (4·9), (4·10), (4·18) and (4·19) into (4·20), (4·20) becomes

$$(\phi^{(2)}_c\,.\,\phi^{(2)}_a) = P^*_L(E^{(0)}_c - E^{(0)}_a)\,(S^{(2)}\phi^{(0)}_c\,.\,\phi^{(0)}_a) + P_L(E^{(0)}_a - E^{(0)}_c)\,(\phi^{(0)}_c\,.\,S^{(2)}\phi^{(0)}_a)$$
$$+ \sum_{\substack{b\,\epsilon\,\deg^{(0)} c\,.\,\deg^{(0)} a\\ E^{(0)}_b \neq E^{(0)}_c,\, E^{(0)}_a}} P^*_L(E^{(0)}_c - E^{(0)}_b)\,(S^{(2)}\phi^{(0)}_c\,.\,\phi^{(0)}_b)\, P_L(E^{(0)}_a - E^{(0)}_b)\,(\phi^{(0)}_b\,.\,S^{(2)}\phi^{(0)}_a)$$
$$(c\,\epsilon\,\deg^{(0)} a,\ E^{(0)}_c \neq E^{(0)}_a). \quad (4{\cdot}21)$$

The set of states, denoted by $\deg^{(0)} c\,.\,\deg^{(0)} a$, comprises the states in common to the two sets of $\deg^{(0)} c$ and $\deg^{(0)} a$. The sum occurring in (4·21), after being transformed into an integral by (4·8), can be evaluated by the following formula, which is valid for any smooth function f (for the establishing of this formula, see later, after (5·15)):

$$\int^{\min.\,(E^{(0)}_a + \frac{1}{2}\Delta E^{(0)}_a,\ E^{(0)}_c + \frac{1}{2}\Delta E^{(0)}_c)}_{\max.\,(E^{(0)}_a - \frac{1}{2}\Delta E^{(0)}_a,\ E^{(0)}_c - \frac{1}{2}\Delta E^{(0)}_c)} P^*_L(E^{(0)}_c - E^{(0)}_b)\, P_L(E^{(0)}_a - E^{(0)}_b) f(E^{(0)}_b)\, dE^{(0)}_b$$
$$= -i\pi P^*_L(E^{(0)}_c - E^{(0)}_a)\,[f(E^{(0)}_a) + f(E^{(0)}_c)]. \quad (4{\cdot}21a)$$

Thus (4·21) becomes

$$(\phi^{(2)}_c\,.\,\phi^{(2)}_a) = P^*_L(E^{(0)}_c - E^{(0)}_a)\Big[(S^{(2)}\phi^{(0)}_c\,.\,\phi^{(0)}_a) - (\phi^{(0)}_c\,.\,S^{(2)}\phi^{(0)}_a)$$
$$- i\pi \sum_{B\neq A} (S^{(2)}\phi^{(0)}_c\,.\,\phi^{(0)}_B(E^{(0)}_a))\,\rho^{(0)}_B(E^{(0)}_a)\,(\phi^{(0)}_B(E^{(0)}_a)\,.\,S^{(2)}\phi^{(0)}_A(E^{(0)}_a))$$
$$- i\pi \sum_{B\neq C} (S^{(2)}\phi^{(0)}_C(E^{(0)}_c)\,.\,\phi^{(0)}_B(E^{(0)}_c))\,\rho^{(0)}_B(E^{(0)}_c)\,(\phi^{(0)}_B(E^{(0)}_c)\,.\,S^{(2)}\phi^{(0)}_a)\Big]$$
$$(c\,\epsilon\,\deg^{(0)} a,\ E^{(0)}_c \neq E^{(0)}_a). \quad (4{\cdot}22)$$

Now, by (4·5), $(\phi_C^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_A^{(0)}(E_a^{(0)}))$ is self-adjoint. The adjoint equation of (4·13) can then be written in the form

$$\begin{aligned}(S^{(2)}\phi_C^{(0)}(E_a^{(0)})\,.\,\phi_A^{(0)}(E_a^{(0)})) &= (\phi_C^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_A^{(0)}(E_a^{(0)}))\\ &\quad+i\pi\sum_{B\neq C,A}(S^{(2)}\phi_C^{(0)}(E_a^{(0)})\,.\,\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_A^{(0)}(E_a^{(0)}))\quad (C\neq A).\end{aligned}\tag{4·23}$$

From (4·23) and (4·13) follows

$$\begin{aligned}&(S^{(2)}\phi_C^{(0)}(E_a^{(0)})\,.\,\phi_A^{(0)}(E_a^{(0)}))-(\phi_C^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))\\ &\quad= i\pi\sum_{B\neq C,A}\{(S^{(2)}\phi_C^{(0)}(E_a^{(0)})\,.\,\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_A^{(0)}(E_a^{(0)}))\\ &\qquad\qquad+(\phi_C^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))\}\\ &\quad= 2i\pi\sum_{B\neq C,A}(S^{(2)}\phi_C^{(0)}(E_a^{(0)})\,.\,\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)})),\quad (C\neq A),\end{aligned}\tag{4·24}$$

(the last step being effected by using (4·13) and (4·23) again, with B for C). Since, for large values of L, the allowance $\varDelta E_a^{(0)}$ can be made arbitrarily small, (4·24) shows that the right-hand side of (4·22) vanishes.

From (4·9), (4·10), (4·18) and (4·19) follows also

$$\begin{aligned}(\phi_a^{(2)}\,.\,\phi_a^{(2)}) &= 1+\sum_{b\,\epsilon\,\deg^{(0)}a}^{E_b^{(0)}\neq E_a^{(0)}}|\,P_L(E_a^{(0)}-E_b^{(0)})\,|^2\,|\,(\phi_b^{(0)}\,.\,S^{(2)}\phi_a^{(0)})\,|^2\\ &= 1+\int_{E_a^{(0)}-\frac{1}{2}\varDelta E_a^{(0)}}^{E_a^{(0)}+\frac{1}{2}\varDelta E_a^{(0)}}|\,P_L(E_a^{(0)}-E_b^{(0)})\,|^2\sum_B|\,(\phi_B^{(0)}(E_b^{(0)})\,.\,S^{(2)}\phi_a^{(0)})\,|^2\,\rho_B^{(0)}(E_b^{(0)})\,dE_b^{(0)},\end{aligned}\tag{4·25}$$

the last step being effected by using (4·8). It will be shown later, (5·15)–(5·17), that, for large values of L, the integral of (4·25) can be neglected in comparison with unity.

If the group $\deg^{(0)}a$ contains no other state $\phi_b^{(0)}$ for which $(\phi_b^{(0)}\,.\,K^{(2)}\phi_a^{(0)})\neq 0$ than the state $\phi_a^{(0)}$ itself, the solution of the secular problem (4·2) is then trivial. Namely,

$$\left.\begin{aligned}&(\phi_a^{(0)}\,.\,\phi_a^{(2)})=1,\quad \text{i.e.}\quad \phi_a^{(2)}=\phi_a^{(0)};\\ \text{also}\qquad &E_a^{(2)}=(\phi_a^{(0)}\,.\,K^{(2)}\phi_a^{(0)}).\end{aligned}\right\}\tag{4·26}$$

For the calculation of the energy values by (4·15) or (4·26), the diagonal matrix element of $K^{(2)}$ can be evaluated as follows. Consider the numerator of the fraction term of (4·3). For an intermediate state $\phi_{a'}^{(2)}$ which is derived from the group $\deg^{(0)}a'$ of states $\phi_{b'}^{(0)}$, of the almost continuous spectrum, one has, by (4·9) and (4·10), and applying (4·8) and (4·12),

$$\begin{aligned}&\sum_{b'\,\epsilon\,\deg^{(0)}a'}(\phi_c^{(0)}\,.\,H^1\phi_{b'}^{(0)})\,(\phi_{b'}^{(0)}\,.\,\phi_{a'}^{(2)})\\ &\quad=(\phi_c^{(0)}\,.\,H^1\phi_{a'}^{(0)})+\sum_{b'\,\epsilon\,\deg^{(0)}a'}^{b'\neq a'}(\phi_c^{(0)}\,.\,H^1\phi_{b'}^{(0)})\,P_L(E_{a'}^{(0)}-E_{b'}^{(0)})\,(\phi_{b'}^{(0)}\,.\,S^{(2)}\phi_{a'}^{(0)})\\ &\quad=(\phi_c^{(0)}\,.\,H^1\phi_{a'}^{(0)})-i\pi\sum_{B'\neq A'}(\phi_c^{(0)}\,.\,H^1\phi_{B'}^{(0)}(E_{a'}^{(0)}))\,\rho_{B'}^{(0)}(E_{a'}^{(0)})\,(\phi_{B'}^{(0)}(E_{a'}^{(0)})\,.\,S^{(2)}\phi_{A'}^{(0)}(E_{a'}^{(0)})).\end{aligned}\tag{4·27}$$

By (4·27) and also its conjugate complex, (4·3) becomes

$$(\phi_c^{(0)}\,.\,K^{(2)}\phi_c^{(0)}) = E_c^{(0)} + (\phi_c^{(0)}\,.\,H^2\phi_c^{(0)}) + \sum_{b'} \frac{(\phi_c^{(0)}\,.\,H^1\phi_{b'}^{(0)})\,(\phi_{b'}^{(0)}\,.\,H^1\phi_c^{(0)})}{E_c^{(0)} - E_{b'}^{(0)}}$$
$$+ \sum_{a' \bar{\epsilon}\, \mathrm{deg}^{(0)} c} \frac{\left|\,(\phi_c^{(0)}\,.\,H^1\phi_{a'}^{(0)}) - i\pi \sum\limits_{B' \neq A'} (\phi_c^{(0)}\,.\,H^1\phi_{B'}^{(0)}(E_{a'}^{(0)}))\,\rho_{B'}^{(0)}(E_{a'}^{(0)})\,(\phi_{B'}^{(0)}(E_{a'}^{(0)})\,.\,S^{(2)}\phi_{A'}^{(0)}(E_{a'}^{(0)}))\,\right|^2}{E_c^{(0)} - E_{a'}^{(0)}}. \tag{4·28}$$

Here the first summation contains only a few intermediate states (arising from $\phi_c^{(0)}$ by virtual absorption of some quanta), while the second summation extends over a large number of intermediate states (arising from $\phi^{(0)}$ by virtual emission of some quanta). The first sum is of the order of L^{-3} and, for large values of L, can be neglected in comparison with the second sum which is independent of L.

5. PHYSICAL INTERPRETATION OF THE SECULAR CHANGE OF BASIS FOR THE STATES OF THE ALMOST CONTINUOUS SPECTRUM

Consider the eigenfunction $\phi_a^{(2)}$ of the secular problem. By (3·15), (4·1), (4·9) and (4·10), it reads

$$\phi_a^{(2)} = \phi_a^{(0)} + \sum_{b \in \mathrm{deg}^{(0)} a}^{E_b^{(0)} \neq E_a^{(0)}} \phi_b^{(0)}\,P_L(E_a^{(0)} - E_b^{(0)})\,(\phi_a^{(0)}\,.\,S^{(2)}\phi_b^{(0)}). \tag{5·1}$$

The physical interpretation of this particular superposition of the states $\phi_b^{(0)}$ of the almost continuous spectrum ($b \in \mathrm{deg}^{(0)} a$) can best be explained in the language of wave mechanics. In wave mechanics, each state $\phi_b^{(0)}$ of the idealized system corresponds to a wave propagation in the configuration space of the quanta, the frequency of the wave propagation being given by $E_b^{(0)}/h$. The superposition (5·1) then represents a *pencil* of wave propagations, which is composed mainly of an *incident* wave $\phi_a^{(0)}$ supported by the appropriate *scattered* waves $\phi_b^{(0)}$ of approximately the same frequency, the amplitude $(\phi_b^{(0)}\,.\,S^{(2)}\phi_a^{(0)})$ $(b \neq a)$ of each scattered wave being determined by the interaction $(\phi_c^{(0)}\,.\,K^{(2)}\phi_b^{(0)})$ $(c \neq b)$, among all the waves in accordance with (4·13). This interpretation is suggested, in the special case where all the states $\phi_b^{(0)}$ considered correspond to only two particles in the fundamental cube, by a comparison with the familiar wave-mechanical theory of atomic collisions. For example, let the rest mass of one of the particles be practically infinite (fixed centre), and let the sides L of the fundamental cube tend to infinity. Because of the interaction, the stationary-state solution $\phi_a^{(2)}$ is of the form (cf. Mott & Massey 1933, chap. II),

$$\phi_a^{(2)} = e^{i\mathbf{k}_a \cdot \mathbf{r}} + r^{-1} e^{ikr} S^{(2)}(\vartheta, \phi) \tag{5·2}$$

for large values of the distance r from the fixed centre.† In this particular case, Dirac (1935, p. 199) has shown, by a direct Fourier transformation, that the scattered wave $r^{-1}e^{ikr}S^{(2)}(\vartheta,\phi)$ can be resolved as a superposition of wave propagations

---

† This holds only for the regular case. In the case of Coulomb interaction, the phase of the incident wave has to be modified according to Gordon (cf. Mott & Massey 1933, chap. III). By a Fourier transformation, it is easily seen that such complication due to the long range of the Coulomb interaction is connected with Fermi's longitudinal quanta of arbitrarily small energy and momentum. As mentioned in the beginning of the last section, such complication will not be considered in this paper.

$$r^{-1}e^{ikr}S^{(2)}(\vartheta,\phi) = \sum_{b\,\epsilon\,\mathrm{deg}^{(0)}a} e^{i\mathbf{k}_b\cdot\mathbf{r}}\left\{\frac{1}{E_a^{(0)}-E_b^{(0)}} - i\pi\delta(E_a^{(0)}-E_b^{(0)})\right\}(\phi_b^{(0)}\,.\,S^{(2)}\phi_a^{(0)}). \qquad (5\cdot3)$$

Here the particular combination of Dirac's delta function and the principle-valued reciprocal function can be regarded as the 'limit' of $P_L(E_a^{(0)}-E_b^{(0)})$ as $L\to\infty$, as seen from (4·11) and (4·12).

In order to show that the above interpretation is valid in the general case, it is convenient to work in the time representation. By (4·1) and (3·37), the secular problem (3·42) reads, for $r = 2$,

$$-\frac{\hbar}{i}\frac{d(\Phi_c^{(0)}\,.\,\Phi_a^{(2)})}{dt} = (\phi_c^{(0)}\,.\,K^{(2)2}\phi_c^{(0)})\,(\Phi_c^{(0)}\,.\,\Phi_a^{(0)})$$
$$+\sum_{b\,\epsilon\,\mathrm{deg}^{(0)}a}^{b\neq c}\exp\left[i(E_c^{(0)}-E_b^{(0)})\,t/\hbar\right](\phi_c^{(0)}\,.\,K^{(2)}\phi_b^{(0)})\,(\Phi_b^{(0)}\,.\,\Phi_a^{(2)}), \qquad (5\cdot4)$$

where, by (3·2), $K^{(2)2} = K^{(2)}-H^0$.

The asymptotic solution which is equivalent to that given by (4·9) and (4·10) now assumes the form

$$(\Phi_a^{(0)}\,.\,\Phi_a^{(2)}) = \exp\left[i(E_a^{(0)}-E_a^{(2)})\,t/\hbar\right],\quad (\Phi_b^{(0)}\,.\,\Phi_a^{(2)}) = 0\quad (E_b^{(0)} = E_a^{(0)}), \qquad (5\cdot5)$$

$$(\Phi_b^{(0)}\,.\,\Phi_a^{(2)}) = \exp\left[i(E_a^{(0)}-E_a^{(2)})\,t/\hbar\right]\frac{\exp\left[i(E_b^{(0)}-E_a^{(0)})\,t/\hbar\right]-1}{E_a^{(0)}-E_b^{(0)}}\,(\phi_b^{(0)}\,.\,S^{(2)}\phi_a^{(0)})$$
$$(b\,\epsilon\,\mathrm{deg}^{(0)}a,\ E_b^{(0)}\neq E_a^{(0)}), \qquad (5\cdot6)$$

which are valid only for the following range of values for t:

$$1/\gamma_a^{(2)} \gg t \gg \hbar/\Delta E_a^{(0)}. \qquad (5\cdot7)$$

Here, the lower bound for t is of the order of L/c, and the condition $t \gg L/c$ is of course necessary in order that the scattered wave reaches its stationary state. The upper bound for t is necessitated by the decay of the amplitude of the incident wave, the decay constant $\gamma_a^{(2)}$ being related to the imaginary part of $E_a^{(2)}$ as follows (compare (5·5)):

$$\gamma_a^{(2)} = -\frac{2}{\hbar}I(E_a^{(2)})$$
$$= \frac{2\pi}{\hbar}R\left\{\sum_{B\neq A}(\phi_A^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))\right\} \quad \text{by (4·15)},$$
$$= \frac{2\pi}{\hbar}\sum_{B\neq A}\rho_B^{(0)}(E_a^{(0)})\,|\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))\,|^2 \quad \text{by (4·23)} \qquad (5\cdot8)$$

Since $\rho_B^{(0)}(E_a^{(0)})$ contains a factor L^3, and, for $B\neq A$, $|\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))\,|^2$ contains a factor L^{-6}, (5·8) shows that $\gamma_a^{(2)}$ contains a factor L^{-3}. The upper bound for t contains therefore a factor L^3. For large values of L, (5·7) can be satisfied without difficulty (say, by letting t contain a factor L^2).

The solution (5·5) and (5·6) in the time representation can be verified directly by substituting the solution into (5·4). For $c = a$, (5·4) then becomes (the common factor $\exp[i(E_a^{(0)} - E_a^{(2)})t/\hbar]$ being removed throughout)

$$E_a^{(2)} - E_a^{(0)} = (\phi_a^{(0)} . K^{(2)2}\phi_a^{(0)}) + \sum_{b \in \deg^{(0)} a}^{b \neq a} (\phi_a^{(0)} . K^{(2)}\phi_b^{(0)}) \frac{1 - \exp[i(E_a^{(0)} - E_b^{(0)})t/\hbar]}{E_a^{(0)} - E_b^{(0)}} (\phi_b^{(0)} . S^{(2)}\phi_a^{(0)}). \tag{5·9}$$

The sum can be transformed by (4·8) into an integral, which can then be evaluated by the following formula:

$$\int_{E_a^{(0)} - \frac{1}{2}\Delta E_a^{(0)}}^{E_a^{(0)} + \frac{1}{2}\Delta E_a^{(0)}} \frac{1 - \exp[i(E_a^{(0)} - E_b^{(0)})t/\hbar]}{E_a^{(0)} - E_b^{(0)}} f(E_b^{(0)})\, dE_b^{(0)} = -i\pi f(E_a^{(0)}). \tag{5·10}$$

(5·10) holds for any function $f(E_b^{(0)})$ smooth at $E_a^{(0)}$ and for any arbitrarily small allowance $\Delta E_a^{(0)}$, provided that t satisfies (5·7) and hence

$$t\Delta E_a^{(0)} > |t(E_b^{(0)} - E_a^{(0)})/\hbar| \gg 1 \quad (b \in \deg^{(0)} a,\ E_b^{(0)} \neq E_a^{(0)}).$$

Thus (5·9) gives the same expression as (4·15) for $E_a^{(2)}$. Further, for $E_c^{(0)} \neq E_a^{(0)}$, (5·4) becomes, by differentiating (5·6),

$$\begin{aligned}(E_a^{(2)} - E_a^{(0)}) (\Phi_c^{(0)} . \Phi_a^{(2)}) &+ \exp[i(E_c^{(0)} - E_a^{(2)})t/\hbar] (\phi_c^{(0)} . S^{(2)}\phi_a^{(0)}) \\ &= (\phi_c^{(0)} . K^{(2)2}\phi_c^{(0)}) (\Phi_c^{(0)} . \Phi_a^{(2)}) + \exp[i(E_c^{(0)} - E_a^{(0)})t/\hbar] (\phi_c^{(0)} . K^{(2)}\phi_a^{(0)}) (\Phi_a^{(0)} . \Phi_a^{(2)}) \\ &\quad + \sum_{b \in \deg^{(0)} a}^{b \neq c, a} \exp[i(E_c^{(0)} - E_b^{(0)})t/\hbar] (\phi_c^{(0)} . K^{(2)}\phi_b^{(0)}) (\Phi_b^{(0)} . \Phi_a^{(2)}).\end{aligned} \tag{5·11}$$

To be consistent with the approximation made above (following (3·17)) in the derivation of the matrix elements for the secular perturbation, and by using (4·1), $(E_a^{(2)} - E_c^{(2)}) (\Phi_c^{(0)} . \Phi_a^{(2)})$ is to be approximated by $(E_a^{(0)} - E_c^{(0)}) (\Phi_c^{(0)} . \Phi_a^{(2)})$. Hence

$$(E_a^{(2)} - E_a^{(0)}) (\Phi_c^{(0)} . \Phi_a^{(2)}) = (E_c^{(2)} - E_c^{(0)}) (\Phi_c^{(0)} . \Phi_a^{(2)}) = (\phi_c^{(0)} . K^{(2)2}\phi_c^{(0)}) (\Phi_c^{(0)} . \Phi_a^{(2)}) \quad (E_c^{(0)} \neq E_a^{(0)}), \tag{5·12}$$

the last step being effected by using (4·15), seeing that, for $E_c^{(0)} \neq E_a^{(0)}$, $(\Phi_c^{(0)} . \Phi_a^{(2)})$ contains the factor $(\phi_c^{(0)} . S^{(2)}\phi_a^{(0)})$ which is already of the order of L^{-3}. Now subtract (5·12) from (5·11), introduce (5·5) and (5·6), and remove the common factor $\exp[i(E_c^{(0)} - E_a^{(0)})t/\hbar]$ throughout. By using (4·8) and (5·10), the result

$$(\phi_c^{(0)} . S^{(2)}\phi_a^{(0)}) = (\phi_c^{(0)} . K^{(2)}\phi_a^{(0)}) + \sum_{\substack{b \in \deg^{(0)} a \\ E_b^{(0)} \neq E_a^{(0)}}}^{b \neq c} (\phi_c^{(0)} . K^{(2)}\phi_b^{(0)}) \frac{1 - \exp[i(E_a^{(0)} - E_b^{(0)})t/\hbar]}{E_a^{(0)} - E_b^{(0)}} (\phi_b^{(0)} . S^{(2)}\phi_a^{(0)})$$

$$(c \in \deg^{(0)} a,\ E_c^{(0)} \neq E_a^{(0)}), \tag{5·13}$$

is easily seen to be the same as (4·13). Thus the solution (5·5) and (5·6) is verified for the range (5·7) of values of t.

The solution (5·5) and (5·6) are connected with the solution (4·9) and (4·10) by (3·39), that is,

$$(\Phi_c^{(0)} . \Phi_a^{(2)}) = \exp[i(E_c^{(0)} - E_a^{(2)})t/\hbar] (\phi_c^{(0)} . \phi_a^{(2)}). \tag{5·14}$$

Hence, for a value $t = t_L$ which lies within the range (5·7)

$$P_L(E_a^{(0)} - E_b^{(0)}) = \frac{1 - \exp[i(E_a^{(0)} - E_b^{(0)})\, t_L/\hbar]}{E_a^{(0)} - E_b^{(0)}}. \tag{5·15}$$

With the help of (5·15) one easily obtains the formula (4·21*a*) by partial fractions. Further, the integral of (4·25) can be evaluated by the formula

$$\int_{E_a^{(0)} - \frac{1}{2}\Delta E_a^{(0)}}^{E_a^{(0)} + \frac{1}{2}\Delta E_a^{(0)}} \left| \frac{1 - \exp[i(E_a^{(0)} - E_b^{(0)})\, t_L/\hbar]}{E_a^{(0)} - E_b^{(0)}} \right|^2 f(E_b^{(0)})\, dE_b^{(0)} = \frac{2\pi t_L}{\hbar} f(E_a^{(0)}), \tag{5·16}$$

which is valid for any smooth function f. By (5·8) and (5·7), the normalization condition (4·25) of the $\phi_a^{(2)}$ is easily verified for large values of L,

$$(\phi_a^{(2)} \,.\, \phi_a^{(2)}) = 1 + \gamma_a^{(2)} t_L \to 1. \tag{5·17}$$

Strictly speaking, for finite values of L, $\Phi_a^{(2)}$ represents a quasi-stationary state of the secular problem. It can be represented by a probability distribution over the stationary states $\Phi_b^{(0)}$ of the idealized system, namely, by (5·5) and (5·6),

$$|(\Phi_a^{(0)} \,.\, \Phi_a^{(2)})|^2 = \exp[-\gamma_a^{(2)} t] = 1 - \gamma_a^{(2)} t, \tag{5·18}$$

$$\begin{aligned} \sum_{b \in \deg_B^{(0)} a} |(\Phi_b^{(0)} \,.\, \Phi_a^{(2)})|^2 &= \int_{E_a^{(0)} - \frac{1}{2}\Delta E_a^{(0)}}^{E_a^{(0)} + \frac{1}{2}\Delta E_a^{(0)}} \left| \frac{1 - \exp[i(E_a^{(0)} - E_b^{(0)})\, t/\hbar]}{E_a^{(0)} - E_b^{(0)}} \right|^2 \\ &\qquad \times |(\phi_B^{(0)}(E_b^{(0)}) \,.\, S^{(2)} \phi_A^{(0)}(E_a^{(0)})|^2 \rho_B^{(0)}(E_b^{(0)})\, dE_b^{(0)} \\ &= \frac{2\pi t}{\hbar} \rho_B^{(0)}(E_a^{(0)})\, |(\phi_B^{(0)}(E_a^{(0)}) \,.\, S^{(2)} \phi_A^{(0)}(E_a^{(0)})|^2 \quad (B \neq A), \end{aligned} \tag{5·18a}$$

which are valid only for the range (5·7) of values of t. As indicated by $\deg_B^{(0)} a$, (5·18*a*) gives the total probability over all states $\Phi_b^{(0)}$ within an arbitrarily small allowance $\Delta E_a^{(0)}$ around $E_a^{(0)}$ which all conform to the detailed specification B. By differentiation with respect to t, it is seen that the rate of decay (namely, $\gamma_a^{(2)}$) of $|(\Phi_a^{(0)} \,.\, \Phi_a^{(2)})|^2$ equals, by (5·8), the sum of the rate of growth (say $\gamma_{BA}^{(2)}(E_a^{(0)})$) for all groups $\deg_B^{(0)} a$,

$$\gamma_{BA}^{(2)}(E_a^{(0)}) = \frac{2\pi}{\hbar} \rho_B^{(0)}(E_a^{(0)})\, |(\phi_B^{(0)}(E_a^{(0)}) \,.\, S^{(2)} \phi_A^{(0)}(E_a^{(0)}))|^2 \quad (B \neq A). \tag{5·19}$$

Similar analysis can be made for the quasi-stationary state $\Psi_a^{(2)}$ of the actual system, viz. by (3·37) and (3·3),

$$\Psi_a^{(2)} = \exp[-iE_a^{(2)} t/\hbar] \left\{ \phi_a^{(2)} + \sum_c \phi_c^{(2)} (\phi_c^{(2)} \,.\, \psi_a^{(2)1}) + \sum_c \phi_c^{(2)} (\phi_c^{(2)} \,.\, \psi_a^{(2)2}) \right\}. \tag{5·20}$$

The probability distribution over the stationary states of the idealized system now extends to states of quite different energy values, $c \mathrel{\bar{\in}} \deg a$ (see (3·22) and (3·27)), which, however, does not grow with time. For $c \in \deg^{(0)} a$, (3·29) shows, with the help of (3·22), that $(\phi_a^{(2)} \,.\, \psi_a^{(2)2})$ is independent of L,

$$(\psi_a^{(2)2} \,.\, \phi_a^{(2)}) + (\phi_a^{(2)} \,.\, \psi_a^{(2)2}) = -\sum_{b \in \deg^{(0)} a} \frac{(\phi_a^{(2)} \,.\, H^1 \phi_b^{(2)})\, (\phi_b^{(2)} \,.\, H^1 \phi_a^{(2)})}{(E_a^{(0)} - E_b^{(0)})^2}, \tag{5·21}$$

while $(\phi_c^{(2)} . \psi_a^{(2)2})$ for $c \in \deg^{(0)} a$ and $c \neq a$ is of the order of L^{-3}. (This distinction is similar to what has been discussed above in connexion with the diagonal *versus* the non-diagonal matrix elements of $K^{(2)}$.) Hence for large values of L,

$$| \Phi_c^{(0)} . \Psi_a^{(2)}) |^2 = \left\{ 1 - \sum_{b \in \deg^{(0)} a} \frac{(\phi_a^{(2)} . H^1 \phi_b^{(2)}) (\phi_b^{(2)} . H^1 \phi_a^{(2)})}{(E_a^{(0)} - E_b^{(0)})^2} \right\} | \Phi_c^{(0)} . \Phi_a^{(0)}) |^2 \quad (c \in \deg^{(0)} a), \tag{5·22}$$

which shows that the variation in time of the relative probabilities $|(\Phi_c^{(0)} . \Psi_a^{(2)})|^2$ for $c \in \deg^{(0)} a$ is the same as that of $|(\Phi_c^{(0)} . \Phi_a^{(2)})|^2$. Now, on the one hand, from the theoretical point of view, one may regard the whole field (including the interactions) within the fundamental cube as a mechanical system, the state of affairs within which is described by the quasi-stationary state $\Psi_a^{(2)}$. On the other hand, from the experimental point of view, the fundamental cube represents roughly the region of space where different elementary particles collide, while outside this region the interaction may be neglected and so the free motion of these particles are described by the stationary states $\Phi_c^{(0)}$. The above analysis of the variation in time of the probability distribution $|(\Phi_c^{(0)} . \Psi_a^{(2)})|^2$ shows that $\Phi_a^{(2)}$ may be used instead of $\Psi_a^{(2)}$ for the description of such collision problems. This justifies the physical interpretation of the secular eigenfunctions $\phi_a^{(2)}$ stated in the beginning of this section. The $\gamma_{BA}^{(2)}(E_a^{(0)})$ of (5·19) thus gives the transition probabilities per unit time for the transition from $\Phi_a^{(0)}$ to the group $\deg_B^{(0)} a$. The transition probabilities can be converted to the cross-sections in the usual way, which will then be independent of L.

In practice, the calculation of the cross-sections for collision problems involves the solution of the system of integral equations (4·13). These integral equations have already been obtained independently by Heitler (1941; also Heitler & Peng 1942*a*) and Wilson (1941), but their derivation did not claim to be mathematically rigorous, nor was the divergence difficulty removed. The present treatment differs from these preliminary treatments by the secular change of basis which enables one to treat the higher-order approximations. Heitler (1941) has also shown, by a comparison with classical electrodynamics in the special case of the non-relativistic scattering of light by an electron, that the cross-sections calculated by (5·19) and (4·13) includes the effect due to the reaction of the radiation (to the first order only), while that calculated by the usual expansion method corresponds to the neglect of this effect. This finding leads to a convenient interpretation of the secular change of basis for the states of almost *continuous spectrum* in the present treatment of collision problems by the method of secular perturbation. Namely, by the secular change of basis (which, mathematically speaking, has to be done), one takes account of the reaction of the radiation. The method of secular perturbation can be continued to higher approximations, involving successive secular change of basis. By so doing, one takes account of the radiation reaction of higher order. Similar interpretation can also be given to the secular change of basis of the states of the almost continuous spectrum where such states act only as intermediate states.

6. THE EFFECTS OF THE RADIATION REACTION—FINITE SELF-ENERGIES AND CROSS-SECTIONS

It will be shown in this section that the divergence difficulty may automatically disappear if the interaction is treated by the method of secular perturbation. For simplicity, this will be demonstrated by means of an example, which deals with the interaction between the meson field and the nucleon field. (A nucleon is either a neutron or a proton.)

Suppose the total charge of the system is $+e$, and the total momentum of the system vanishes. Let $\phi_P^{(0)}$ denote the state which corresponds to a proton-like quantum at rest in the fundamental cube; let $\phi_a^{(0)}$ denote the state which corresponds to a meson-like quantum of charge $+e$ and momentum $\mathbf{p}_a$, together with a neutron-like quantum of equal but opposite momentum. The energy values of these states are

$$E_P^{(0)} = m_P c^2, \quad E_a^{(0)} = \sqrt{(m_N^2 c^4 + c^2 p_a^2)} + \sqrt{(m_M^2 c^4 + c^2 p_a^2)}, \tag{6·1}$$

where m_P, m_N and m_M denote the rest masses of the proton-like, neutron-like, and meson-like quanta respectively. For simplicity, assume $m_P = m_N$ which is practically infinite, and neglect the dependence of the matrix elements $(\phi_P^{(0)} . H^1 \phi_a^{(0)})$ on the spin orientations and the directions of the momenta of the quanta, so

$$\left.\begin{aligned} E_a^{(0)} - E_P^{(0)} &= \sqrt{(m_M^2 c^4 + c^2 p_a^2)} \equiv \epsilon_a, \text{ say,} \\ (\phi_P^{(0)} . H^1 \phi_a^{(0)}) &\equiv H(\epsilon_a), \text{ say,} \end{aligned}\right\} \tag{6·2}$$

which all depend only on the energy ϵ_a of the meson-like quantum. With this simplification, the kernel of the integral equation (4·13) reduces to a constant,

$$\begin{aligned} (\phi_C^{(0)}(E_a^{(0)}) . K^{(2)} \phi_B^{(0)}(E_a^{(0)})) &= \frac{(\phi_C^{(0)}(E_a^{(0)}) . H^1 \phi_P^{(0)})\,(\phi_P^{(0)} . H^1 \phi_B^{(0)}(E_a^{(0)}))}{E_a^{(0)} - E_P^{(0)}} \\ &= \frac{|H(\epsilon_a)|^2}{\epsilon_a} \equiv K(\epsilon_a), \text{ say,} \quad (C \neq B). \end{aligned} \tag{6·3}$$

The density $\rho_B^{(0)}(E_a^{(0)})$ of states is given by

$$\rho_B^{(0)}(E_a^{(0)}) = L^3 c^{-2} h^{-3} p_a \epsilon_a d\Omega_B, \tag{6·4}$$

where $d\Omega_B$ denotes an element of solid angle around the direction of the momentum of, say, the meson-like quantum. The solution of (4·13) is then trivial,

$$(\phi_B^{(0)}(E_a^{(0)}) . S^{(2)} \phi_A^{(0)}(E_a^{(0)})) = \frac{K(\epsilon_a)}{1 + i\pi K(\epsilon_a) \rho(\epsilon_a)} \equiv S(\epsilon_a), \text{ say} \quad (B \neq A), \tag{6·5}$$

with

$$\rho(\epsilon_a) \equiv \sum_B \rho_B^{(0)}(E_a^{(0)}) = L^3 c^{-2} h^{-3} p_a \epsilon_a 4\pi. \tag{6·6}$$

The transition probability (5·19) for the scattering of mesons of energy ϵ_a by a nucleon is, by (6·5),

$$\gamma_{BA}^{(2)}(E_a^{(0)}) = 2\pi \rho_B^{(0)}(E_a^{(0)}) \left| \frac{K(\epsilon_a)}{1 + i\pi K(\epsilon_a) \rho(\epsilon_a)} \right|^2 \quad (B \neq A), \tag{6·7}$$

which differs from the usual expression given by the expansion method, namely,

$$2\pi\rho_B^{(0)}(E_a^{(0)})\,|\,(\phi_B^{(0)}(E_a^{(0)})\,.\,K^{(2)}\phi_A^{(0)}(E_a^{(0)}))\,|^2 = 2\pi\rho_B^{(0)}(E_a^{(0)})\,[K(\epsilon_a)]^2 \quad (B \neq A), \tag{6·8}$$

by the factor $\{1+[\pi K(\epsilon_a)\,\rho(\epsilon_a)]^2\}^{-1}$. In the usual meson theory, for large values of ϵ_a, one has ($\propto$ meaning 'varies as'; g denoting the interaction constant)

$$H(\epsilon_a) \propto L^{-\frac{3}{2}} g \epsilon_a^{\frac{1}{2}}, \quad \text{so} \quad K(\epsilon_a)\,\rho(\epsilon_a) \propto L^0 g^2 \epsilon_a^2 \tag{6·9}$$

by (6·3) and (6·6). The effect of the radiation reaction, as indicated by the appearance of the factor $\{1+[\pi K(\epsilon_a)\,\rho(\epsilon_a)]^2\}^{-1}$ in (6·7), is thus a reduction of the cross-sections for the scattering of mesons at high energies of the mesons. (For more accurate calculations of the scattering cross-sections, see Ma (1943), Heitler & Peng (1943), Ma & Hsüeh (1944); but the results are qualitatively similar to (6·7) discussed above. The effect of the radiation reaction makes the cross-sections a few hundred times smaller and is favoured by comparison with measurements on the scattering of cosmic-ray mesons by atomic nuclei (Heitler & Peng 1942*b*).

Consider now the self-energy of a proton, which, for the second approximation, is given by (cf. (4·28))

$$E_P^{(2)} - E_P^{(0)} = \sum_a \frac{\left|(\phi_P^{(0)}\,.\,H^1\phi_a^{(0)}) - i\pi \sum\limits_{B\neq A} (\phi_P^{(0)}\,.\,H^1\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))\right|^2}{E_P^{(0)} - E_a^{(0)}}. \tag{6·10}$$

From (6·2), (6·5) and (6·6) follows

$$\begin{aligned}(\phi_P^{(0)}\,.\,H^1\phi_a^{(0)}) - i\pi \sum_{B\neq A} (\phi_P^{(0)}\,.\,H^1\phi_B^{(0)}(E_a^{(0)}))\,\rho_B^{(0)}(E_a^{(0)})\,(\phi_B^{(0)}(E_a^{(0)})\,.\,S^{(2)}\phi_A^{(0)}(E_a^{(0)}))& \\ = H(\epsilon_a) - i\pi H(\epsilon_a)\,\rho(\epsilon_a)\,S(\epsilon_a) = \frac{H(\epsilon_a)}{1+i\pi K(\epsilon_a)\,\rho(\epsilon_a)}&.\end{aligned} \tag{6·11}$$

Hence (6·10) becomes

$$\begin{aligned} E_P^{(2)} - E_P^{(0)} &= -\sum_a \left|\frac{H(\epsilon_a)}{1+i\pi K(\epsilon_a)\,\rho(\epsilon_a)}\right|^2 \frac{1}{\epsilon_a} \\ &= -\int_{m_M c^2}^{\infty} \frac{|H(\epsilon_a)|^2}{\epsilon_a} \frac{\rho(\epsilon_a)\,d\epsilon_a}{1+[\pi K(\epsilon_a)\,\rho(\epsilon_a)]^2} \quad \text{by (4·8) and (6·7),} \\ &= -\int_{m_M c^2}^{\infty} \frac{K(\epsilon_a)\,\rho(\epsilon_a)}{1+[\pi K(\epsilon_a)\,\rho(\epsilon_a)]^2} d\epsilon_a \quad \text{by (6·4).} \end{aligned} \tag{6·12}$$

This differs from the expression (which is usually taken to be the self-energy in the expansion method)

$$\sum_a \frac{|(\phi_P^{(0)}\,.\,H^1\phi_a^{(0)})|^2}{E_P^{(0)} - E_a^{(0)}} = -\int_{m_M c^2}^{\infty} \frac{|H(\epsilon_a)|^2}{\epsilon_a} \rho(\epsilon_a)\,d\epsilon_a = -\int_{m_M c^2}^{\infty} K(\epsilon_a)\,\rho(\epsilon_a)\,d\epsilon_a, \tag{6·13}$$

again by the factor $\{1+[\pi K(\epsilon_a)\,\rho(\epsilon_a)]^2\}^{-1}$ in the integrand. This factor arises from the secular change of basis of the intermediate states, and may be interpreted as the effect of radiation reaction. For the usual meson theory, by (6·9), the integral (6·12) for the self-energy converges, like $\int^{\infty} \frac{d\epsilon}{\epsilon^2}$, in the method of secular perturbation, while the integral (6·13) diverges, like $\int^{\infty}\epsilon^2 d\epsilon$, in the expansion method. The mathematical origin of the divergence difficulty arises from the fact that the factor $\{1+[\pi K(\epsilon_a)\rho(\epsilon_a)]^2\}^{-1}$, though it assumes the value unity for vanishing interaction ($g=0$), cannot be expanded as a power series in the interaction constant g for *all* values of ϵ_a.

For the calculation of the cross-sections for the scattering of mesons by a nucleon to the fourth approximation, say, the fourth-order secular perturbation $(\phi_c^{(0)}.K^{(4)}\phi_b^{(0)})$ can be derived by applying the expansion method in the basis $\{\phi_b^{(4)}\}$ and then changing the basis, say to $\{\phi_b^{(0)}\}$. (Compare (3·25) and (3·35) for an illustration.) The non-diagonal element of $K^{(4)}$ contains in particular the sum (note that, for the non-degenerate state P, $\phi_P^{(4)}=\phi_P^{(2)}=\phi_P^{(0)}$)

$$\sum_{b'\in\deg^{(2)}b}\frac{(\phi_c^{(0)}.H^1\phi_P^{(0)})\sum\limits_{a'\in\deg^{(0)}b'}(\phi_P^{(0)}.H^1\phi_{a'}^{(0)})(\phi_{a'}^{(0)}.\phi_{b'}^{(4)})\sum\limits_{c'\in\deg^{(2)}b'}(\phi_{b'}^{(4)}.\phi_{c'}^{(0)})(\phi_{c'}^{(0)}.H^1\phi_P^{(0)})(\phi_P^{(0)}.H^1\phi_b^{(0)})}{(E_b^{(2)}-E_P^{(2)})(E_b^{(2)}-E_{b'}^{(2)})(E_b^{(2)}-E_P^{(2)})},$$
$$(c\in\deg^{(2)}b,\ c\neq b). \qquad (6{\cdot}14)$$

In the expansion method, the secular change of basis in the numerator is neglected (also $E^{(0)}$ is used for $E^{(2)}$, but this is not serious), resulting in a divergent expression. In the method of secular perturbation, the non-diagonal elements (6·14) are needed for the determination of the secular eigenfunctions $\phi_a^{(4)}$ for the fourth approximation, and for this purpose, (6·14) may be approximated by

$$\sum_{b'\in\deg^{(2)}b}\frac{(\phi_c^{(0)}.H^1\phi_P^{(0)})\sum\limits_{a'\in\deg^{(0)}b'}(\phi_P^{(0)}.H^1\phi_{a'}^{(0)})(\phi_{a'}^{(0)}.\phi_{b'}^{(2)})\sum\limits_{c'\in\deg^{(0)}b'}(\phi_{b'}^{(2)}.\phi_{c'}^{(0)})(\phi_{c'}^{(0)}.H^1\phi_P^{(0)})(\phi_P^{(0)}.H^1\phi_b^{(0)})}{(E_b^{(2)}-E_P^{(2)})(E_b^{(2)}-E_{b'}^{(2)})(E_b^{(2)}-E_P^{(2)})},$$
$$(c\in\deg^{(2)}b,\ c\neq b). \qquad (6{\cdot}15)$$

The approximation consists of replacing $\phi_{b'}^{(4)}$ by $\phi_{b'}^{(2)}$, which corresponds to that used above from (4·4) to (4·5) in the case of the second approximation. Because of the effect due to the secular change of basis from $\{\phi_{a'}^{(0)}\}$ to $\{\phi_{b'}^{(2)}\}$ for the intermediate states, (6·15), like (6·12), converges. Hence $(\phi_c^{(0)}.K^{(4)}\phi_b^{(0)})$ $(c\in\deg^{(2)}b,\ c\neq b)$, is of bounded variation, and the asymptotic solution of the secular problem can be obtained similarly. The transition probabilities $\gamma_{BA}^{(4)}(E_a^{(2)})$ can be calculated from the corresponding $(\phi_b^{(0)}.S^{(4)}\phi_a^{(0)})$ $(b\in\deg^{(2)}a,\ b\neq a)$, without any divergence difficulty.

The self-energy for the fourth approximation can then be calculated with the help of $(\phi_b^{(0)}.S^{(4)}\phi_a^{(0)})$ $(b\in\deg^{(2)}a,\ b\neq a)$. The effect of the radiation reaction of the second order is thus included.

I wish to express my great indebtedness to Ma & Hsüeh (1944). Without their new derivation of Heitler's and Wilson's integral equations in the energy representation with the help of Dirac's delta function combined with the principle-valued reciprocal function, the above task to link up Heitler's and Wilson's preliminary

theory with the rigorous perturbation theory for a system of almost continuous spectrum would have been more difficult.

I am also very much obliged to Professor Max Born for his continuous encouragement.

REFERENCES

Born, Heisenberg & Jordan 1926 *Z. Phys.* **35**, 557.
Born & Jordan 1930 *Quantenmechanik*. Springer: Berlin.
Born & Peng 1944 *Proc. Roy. Soc. Edinb.* **62**, 127.
Dirac 1927 *Proc. Roy. Soc.* **114**, 243, 710.
Dirac 1935 *Quantum mechanics*, 2nd ed. Oxford.
Fermi 1932 *Rev. Mod. Phys.* **4**, 87.
Heisenberg & Pauli 1929 *Z. Phys.* **56**, 1; **59**, 168.
Heitler 1941 *Proc. Camb. Phil. Soc.* **37**, 291.
Heitler & Peng 1942*a* *Proc. Camb. Phil. Soc.* **38**, 296.
Heitler & Peng 1942*b* *Phys. Rev.* **62**, 81.
Heitler & Peng 1943 *Proc. Roy. Irish Acad.* **49**, 112.
Jordan & Pauli 1928 *Z. Phys.* **47**, 151.
Jordan & Wigner 1928 *Z. Phys.* **47**, 631.
Ma 1943 *Proc. Camb. Phil. Soc.* **39**, 168.
Ma & Hsüeh 1944 *Proc. Camb. Phil. Soc.* **40**, 167.
Mott & Massey 1933 *Atomic collisions*. Oxford.
Pauli 1941 *Rev. Mod. Phys.* **13**, 203.
Schrödinger 1926 *Ann. Phys., Lpz.*, **80**, 437.
Wilson 1941 *Proc. Camb. Phil. Soc.* **37**, 301.

A discussion on Units and Standards

21 *March* 1946

[Plate 6]

	PAGE
Introduction (Sir Charles Darwin, F.R.S.)	149
The standards of length (J. E. Sears)	152
The standards of length in wave-lengths of light (H. Barrell)	164
The standards of mass (F. A. Gould)	171
The temperature scale (J. A. Hall)	179
Electrical standards (L. Hartshorn)	185
The acceleration due to gravity (J. S. Clark)	192
The barometric standard (F. A. Gould)	195
The measurement of volume (V. Stott)	200
The photometric units and standards (Light Division)	204
Acoustical standards of measurement (R. S. Dadson)	207

INTRODUCTION

BY SIR CHARLES DARWIN, F.R.S.

The sequence of papers following, contributed by members of the staff of the National Physical Laboratory, gives an account of most of the units and standards which are studied at the Laboratory. Each paper explains briefly the principal technical points involved, together with a short history. Of necessity a standard is to some extent a convention, and it may therefore be well to introduce the subject by a short review of the legal and conventional authorities on which they depend.

For the metric system the main international authority is the Conférence Générale des Poids et Mesures, at which duly accredited delegates of each of the thirty-two nations subscribing to the 1875 Treaty, the Convention du Mètre, are present, and which, in normal circumstances, meets every six years. The Comité International des Poids et Mesures is the Executive body appointed by the Conférence Générale to manage the affairs of the Bureau International des Poids et Mesures between meetings of the Conférence Générale. It consists of eighteen members, elected by the Conference, each of whom must belong to a different State, but who, as members of the Committee, are no longer representatives only of their individual countries, but of all subscribing countries. The International Committee meets every two years, makes recommendations to the General Conference, or acts on authority delegated to it by the Conference.

There are also three Advisory Committees which deal respectively with electricity, photometry and temperature. These Committees meet whenever necessary and make their recommendations to the International Committee which, if it approves them, offers them to the General Conference for ratification. The original Convention covered length and mass, and matters incidental thereto, such as thermometry and barometry; it was amended in 1921 to bring in the electrical and photometric units. The headquarters of the organization are at the Bureau International at Sèvres.

Although the countries subscribing to the Convention du Mètre look to the Bureau International as the central authority for all measures relating to the metric system, this body has no direct legal authority in any individual country. It may be well, therefore, to touch on the legal situation in Britain. Our standards go back a long way—there are still in existence standards authorized by Henry VII, Henry VIII and Elizabeth—and due regard was paid by successive administrations to the need of properly legalized standards for purposes of trade and industry. From the scientific angle, however, it must be admitted that the legislation has not been at all systematic. The present legal standards of length and mass, the yard and pound (avoirdupois) were remade after the destruction of the old standards, when the Houses of Parliament were burnt down in 1834, and they were equated as nearly as could be ascertained to the old standards. The provisions for their use were embodied in the Act of 1856, but the specifications were incomplete in some respects, and in others they prescribed methods of measurement which were subsequently shown to be capable of improvement. Similar defects are to be found in the legal definition of the unit of cubic capacity.

For the electrical standards, on the other hand, a better policy was adopted, and the relevant Act made no quantitative specifications, but provided that they should be determined by Order in Council. The legal standard of the volt is at present embodied in an instrument at Teddington, which later experience has shown to be somewhat defective; however, when a certification is demanded it is always for the international and not for the legal volt, so that there have been no serious consequences of the defects.

Though many Acts of Parliament refer to temperature 'on Fahrenheit's scale', there is no enactment that prescribes what this scale is; it is now conventionally taken to be the absolute thermodynamic scale with 32° and 212° F exactly equal to 0° and 100° C. This certainly differs appreciably from the Fahrenheit scale as realized by the mercury-in-glass thermometers which were actually employed at the time the standards were laid down. Similarly, in calorimetry various enactments refer to the British thermal unit, but never specify it, and there are no less than three equally good authorities who might be quoted, all giving slightly different values for it. In photometry, candle-power was defined in 1860 in terms of a sperm candle, but it has been changed since then without formal authority, first to a flame lamp and later to electric lamps.

The general outcome of the papers following is summarized in table 1, which calls for some explanation. Apart from the need that it should keep constant in time, the important character of a fundamental standard is its accuracy of comparison, that is to say the degree of precision with which copies can be compared with their prototype. For derived standards there are two differing qualities of this type, accuracy of comparison and accuracy of realization. The first is the relative accuracy with which two examples of the same unit can be compared, the second the evaluation of an example of the unit in terms of the fundamental standards of length, mass and time, and it will be noted that in many cases the second is distinctly inferior in accuracy to the first.

TABLE 1. TABLE OF ACCURACIES

Unit	Physical Dimensions	Standard or Reference	Nature	Accuracy of Realization parts per million	Accuracy of Comparison parts per million
length	L	international metre	material	—	0·2
		wave-length of light	natural	0·02	0·01 (1)
mass	M	international kilogram	material	—	0·002 (2)
time	T	mean solar second	natural	—	0·01
temperature	—	thermodynamic scale	theoretical	— (3)	1 (4)
volume	L^3	litre	derived (5)	1	—
gravity	LT^{-2}	earth's attraction	natural	2	1 (6)
pressure	$ML^{-1}T^{-2}$	standard barometer	—	10	—
electric resistance	LT^{-1}	ohm	theoretical	10	1
electric current	$M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}$	ampere	theoretical	10	1 (7)
quantity of heat	ML^2T^{-2}	calorie	—	100 (8)	—
luminous intensity	ML^2T^{-3}	candle (9)	material	2,000	1,000
sound pressure	$ML^{-1}T^{-2}$	—	derived	10,000	— (10)

(1) Comparison of derived end-standards (metre or yard).
(2) Comparison of kilograms.
(3) The absolute zero is accepted as $-273{\cdot}15 \pm 0{\cdot}02°$ C.
(4) Reproducibility of the ice-point on the absolute scale.
(5) Derived from the kilogram of water; 1000·027 cm.^{3} = 1 l.
(6) Accuracy of relative comparisons between different stations.
(7) The relative comparison is not of amperes but of volts.
(8) Electrical determination of the mechanical equivalent for the 15° C cal.
(9) The dimensions are those of power, but there is an additional factor representing the response of the eye to energy at any wave-length.
(10) The acoustic unit selected is the most generally useful; the accuracy of comparison varies with the experimental conditions but would generally be better than the accuracy of realization.

The succeeding papers deal with all the items in the table with two exceptions. One of these is time, and though much work is done on the subject at the laboratory, it is primarily the responsibility of the Astronomer Royal. We owe to him the information entered in the table, which is roughly speaking the accuracy of behaviour of the best crystal clocks; already these clocks are so good that they can show up small irregularities in the earth's rotation. Time, of course, plays a part in many of the derived standards, but the accuracy of realization of these is for the most part much lower, so that time-keeping is not usually the most serious difficulty in their attainment.

There is also no paper on calorimetry. In modern times precise calorimetry is always done electrically, since the measurement of electrical energy is much easier than that of any other form. Indeed, it has been proposed that when the absolute system of electrical units is adopted in place of the present International System, the calorie shall be redefined as a definite quantity of electric energy by the equation 1 cal. = 3600/860 joules or 1/860 watt-hour. The effect of this reform will be to remove the question of the specific heat of water from the place of importance which it now occupies, and which it does not deserve.

One entry in the table is the value of gravity, which is not strictly speaking a standard at all. Its presence here is justified because it is an essential intermediary in the evaluation of several of the other standards, in particular those of pressure and of electric current.

The very different natures of all these physical quantities impose some strain in forcing them into a single table. In each case there is some range of values where the measurement is easiest and most precise: for example, absolute temperature is very much more precisely known at ordinary temperatures than at a red-heat. In the table the most favourable case has been taken, and the choice is partly explained by footnotes, but for full appreciation the succeeding papers must be studied.

THE STANDARDS OF LENGTH

BY J. E. SEARS

The fundamental standards

The units of length, on both the British and metric systems, are at present defined by material standards known respectively as the Imperial Standard Yard, and the International Prototype Metre.

The Imperial Standard Yard is a bar of Baily's metal (16 parts copper, 2½ parts tin and 1 part zinc) 1 in. square in section and 38 in. long. It is one of a number of similar bars, cast in 1845, and was given legal status as the Imperial Standard by the Weights and Measures Act of 1856. It replaced the former standard—the 'Bird' Yard—which had been destroyed in the fire at the Houses of Parliament in 1834. At 1 in. from either end there is a hole ½ in. diameter and ½ in. deep, the bottom of which therefore lies in the neutral plane of the bar. At the centre of the bottom surface of each hole a polished gold stud is inserted on which three lines are ruled, at right angles to the length of the bar. The length of the yard is defined, in the Weights and Measures Act, 1878, as the distance between the central lines at the two ends when the bar is at the temperature of 62° 'of Fahrenheit's thermometer'.

The International Prototype Metre is a bar of platinum-iridium alloy (90 % platinum, 10 % iridium) of a special winged X-form section devised by G. Tresca to give maximum rigidity in relation to the weight of metal used; the neutral plane of the section is exposed throughout the length of the bar, and the metre is defined as the distance between two transverse graduations on the neutral plane, near the ends of the bar, when the latter is at the temperature of 0° C. This bar is one of a number prepared in pursuance of the International Treaty of 1875, known as the Convention du Mètre, and received sanction as the International Prototype, replacing the old Mètre des Archives, at the first General Conference of Weights and Measures held in pursuance of the Treaty, in 1889.*

The reason for putting the graduations on the neutral plane is that the material of the bar in this plane is neither extended nor compressed by the elastic distortion of the bar due to its own weight when resting on its supports. It is still necessary, however, to specify the manner in which the bar shall be supported. If supported at its two ends, for example, it would droop in the middle, and the distance between the graduation marks would be slightly shorter than the 'free' distance between them. The Weights and Measures Act, 1878, specifies that the Imperial Standard Yard shall be supported on a system of eight rollers connected by pivoted levers in such a manner that the weight of the bar is equally distributed between them, thus minimizing the flexure of the bar, and providing freedom for expansion or contraction according to temperature conditions. The International Prototype Metre is supported for measurement on two rollers symmetrically arranged at a separation of 571 mm., this distance being calculated so that the separation of the graduations is a maximum and therefore, to the first order, independent of slight errors in positioning of the rollers. Actually the difference in the length of the yard as supported on eight rollers, or alternatively as supported on two rollers suitably placed to give maximum length, is quite negligible in relation to the accuracy of measurement, and it is found preferable in practice to make comparisons on two supports only, since there is less risk in this way of accidental movements of the bar during measurement.

Comparisons of line-standards

It will be noted that both the Imperial Standard Yard and the International Prototype Metre bars are 'line' standards, adapted to be compared with other measures by means of micrometer microscopes on a comparator. On the comparator two microscopes are used, set at a fixed distance apart, to read the lines at the two ends of the bar, and the two bars to be compared are carried on two parallel girders on the carriage of the comparator, with independent adjustments for aligning the bars and focusing the graduations in the microscopes. The carriage can be moved so as to bring either bar at will into view in the microscopes. In the eyepiece of each microscope a pair of parallel cross-wires is adjusted by means of the micrometer so as to appear symmetrically disposed in relation to the image of the graduation on the bar, and the reading recorded.

In this way the lengths of two bars can be compared. But owing to the effects of possible asymmetry in the graduations and in the cross-wires of the microscopes, and of the 'personal equations' of setting of the observers, it is necessary to repeat the readings with the bars reversed in direction, and for each observer to take an equal number of readings with each microscope under each condition. Finally, in order to achieve the greatest possible accuracy it is usual to compare not merely two bars, but a series of bars, each with each, and to evaluate from the totality of the measurements, by the method of least squares, the most probable values of the differences between the individual bars.

---

* 1889, *C.R. Première Conférence Générale des Poids et Mesures.* Paris: Gauthier-Villars.

It should be remarked that the graduation marks on a line-standard are, in the last resort, in the nature of furrows, with more or less ragged edges, whose appearance will differ according to the magnifying power of the microscopes employed, and according to the manner in which they are illuminated. Both these factors will affect the observer's estimate of the mean position of the line. The complete specification of a line-standard should therefore, strictly speaking, include instructions on these two points. Generally speaking it is undesirable to employ a magnification exceeding × 100; × 50 is more usual. The exact reproduction of illumination conditions is in any case a matter of some difficulty. The actual nature and quality of the graduations has, of course, a marked bearing on this question, and the metric standards have the advantage over the older British standards in this respect.

Temperature effects

It is necessary, of course, to maintain a close control and make accurate measurements of the temperature of the bars during comparison. For this reason the comparator is arranged so that the bars can be immersed, during measurement, in the inner compartment of a double water-bath, provided with stirring devices to maintain a uniform temperature. In the case of the metric standards it is permissible to immerse the bars in distilled water, together with suitable thermometers, during measurement, but the construction of the British standards makes this procedure undesirable in their case, and the comparisons have to be made in air, so that the control of temperature presents greater difficulty, though the outer water-bath can still be used. In any case measurements cannot readily be made exactly at 62° F, and still less can they conveniently be made at 0° C. Since different bars will have different coefficients of thermal expansion, it is necessary to determine these and make allowances for them in computing the results of comparisons actually made at other temperatures.

The process of determining coefficients of expansions is similar to that of comparing two bars, except that the bars being compared are kept in separate compartments on the comparator (or dilatometer as this variant of the apparatus is named), one of which is kept at a constant temperature while the temperature of the other is varied by known amounts between successive comparisons.

It is important to note that the *units* are absolute measures of spatial extension, and entirely independent of any temperature considerations, once their values have been ascertained in terms of the defining standards under the prescribed conditions. There has in the past been a certain confusion of ideas on this point, with the result, for example, that it was at one time considered that a metre of any material must have the length of a metre at the temperature of definition, 0° C, and that, in consequence, metres of varying materials would be of varying lengths at ordinary temperatures. It has now been agreed internationally that industrial standards of length, whether British or metric, shall be adjusted to be correct—i.e. to agree in value with the units—at 20° C (68° F). So long as such standards are made of materials having a similar coefficient of expansion to the articles they are used to measure, and the standards and the articles are at the *same* temperature at the time of measurement, the comparison will then give directly the true size of the article at 20° C, whatever the actual temperature of measurement.

Materials of construction for length standards

This brings us to a consideration of the kinds of materials used for the construction of standards for various purposes, other than the primary standards. The principal materials so employed are steel, brass and invar, together with fused silica and crystal quartz. Steel and brass are used for industrial standards for the measurement of articles made of materials with corresponding coefficients of expansion. In some cases brass, or gun-metal, gauges are used in preference to steel in explosives factories simply for safety reasons.

Invar, an alloy steel containing 36 % nickel, has the remarkable property of having a very low coefficient of thermal expansion, varying according to the size and shape of the piece and the thermal and mechanical treatment to which it has been subjected, from about $1{\cdot}5 \times 10^{-6}$ to $-0{\cdot}5 \times 10^{-6}/1^\circ$ C. It is, however, subject to a marked secular growth which renders it unsuitable for use as a reference standard, and its low coefficient makes it also unsuitable as a standard of comparison for use in most industrial measurements. Apart from its special applications in the construction of such articles as thermostats, pendulum rods for clocks and the like, it serves two very useful purposes in metrology, namely, it provides a convenient comparison standard for the determination of coefficients of expansion, and is very valuable as a material of construction for the tapes and wires which are used for the measurement of bases in geodetic surveys where errors due to uncertainty of temperature measurements in the field, with materials of higher coefficients, might be of serious consequence.

Fused silica, with a coefficient of expansion of about $0{\cdot}4 \times 10^{-6}/1^\circ$ C, was considered likely to be a stable material, and therefore as possibly helping to provide a control on the constancy of units of length defined by material standards made of alloy metals. Two standards of fused silica have been made, one of which is at the N.P.L. (Kaye 1911) and one with the Indian Ordnance Survey. Comparisons of the silica metre at the N.P.L. since 1911 with other copies of the metre linked back to the platinum-iridium standard have exhibited no systematic change.

Crystal quartz is another material which, by virtue both of its structure and its antiquity, might be expected to have a high degree of stability. Standards of crystal quartz are necessarily limited to a fraction of the length of the normal unit. End-standards of quartz of 1 dm. length have, however, been made and measured with high precision by optical interference methods. This work has not yet been repeated systematically over a long enough period to draw any definite conclusions from it.

Constancy of units

Control on the reproducibility of the units in the event of loss or damage is maintained by the provision of copies, in the case of the Imperial Standards the Parliamentary Copies, which are held in the custody of the Board of Trade, the Royal Observatory, the Royal Mint and the Royal Society, and in the case of the Metric Prototype Standards the 'Témoins', which are held at the Bureau International des Poids et Mesures, and the various national copies distributed to all the countries subscribing to the Convention du Mètre. Periodical intercomparisons between these copies suffice to show whether any one of them has changed relative to the others, and very occasional checks between the principal copies and the fundamental standards maintain the connexion with the latter. Should one of the fundamental standards be found at any time to have changed relative to the copies by an amount appreciably exceeding the experimental error of the comparisons, steps would need to be taken to replace it, or to substitute for it one of the copies presumed not to have changed. It must be remarked, however, that continued agreement between a number of standards all of the same material does not, by itself, suffice to establish the absolute constancy of any of them.

With regard to the continuity of the units and the stability of the standards the following may be stated. There are still in existence, at the South Kensington Museum, the old standard yards of Henry VII and Elizabeth. Both these are end-standards, with roughly incised subdivisions representing fractions of the yard. The former has been found to agree with our present yard within $\frac{1}{32}$ in., and the latter, though broken and roughly mended, within 0·01 in.* This bar, which has since been more carefully repaired, remained the standard of the country until 1824, when it was superseded by a yard known as the Bird Yard, whereon the unit length was defined as the distance between two dots on inserted gold plugs, which had been placed in the custody of the Clerk of the House of Commons in 1760, following the Reports of Lord Carysfort's Committee,† but which did not receive legal sanction till 64 years later (Weights and Measures Act, 1824). Unfortunately, its legal existence was all too short, for it was destroyed, as mentioned above, in 1834.

The Weights and Measures Act, 1824, prescribed that if the Imperial Standard Yard were ever lost or destroyed it should be restored by reference to the length of the pendulum beating seconds in London. In a certain sense, then, the unit of that day was theoretically a 'natural' unit, of which the standard was no more than the material representation. The Committee entrusted with the duty of producing a new standard soon found, however, that the difficulty of reproducing with accuracy the length of the seconds pendulum was greater than that of reproducing the length of the old bar through the agency of other standards which were known to have been compared with it, and this procedure was that which, in fact, was employed for the establishment of the new (and present) standard. It is estimated that the new standard represented the current values of the best scientific standards of the time (which were susceptible of intercomparison with higher accuracy than the Imperial Standard itself) within 0·0002 in. and was in agreement with the standard it replaced within the limit of accuracy to which the latter could be measured.

---

* 1873, *Report of the Warden of the Standards on the Proceedings of the Standard Department the Board of Trade.*

† 1738–65, *Reports from Committees of the House of Commons, Miscellaneous Subjects*, **2**, 434.

Evidence for the constancy of the yard in recent times is fairly satisfactory. The results of the comparisons of the Imperial Standard Yard with its three contemporary Parliamentary Copies, Nos. 2, 3 and 5, in 1912, 1922 and 1932,* show that each of the four bars has remained in the same relationship to the mean of the group as when first determined in 1852, within 0·00005 in. In particular P.C. 2 has remained constant in relation to the mean within 0·000004 in., and P.C. 5 within 0·000016 in. for the last three comparisons. It should be mentioned, however, that a new Parliamentary Copy authorized to be made by the Weights and Measures Act, 1878, was constructed as nearly similar as possible to the original series in 1879. Since its first determination, in 1886, this bar has shown a systematic shortening in relation to the rest of the group, successive comparisons, when plotted, falling very closely on a smooth curve, with an asymptotic value of -228×10^{-6} in. which had very nearly been reached 40 years after its construction. If it is assumed that the original bars behaved similarly during the first 40 years of their existence dating from 1845, then at the date of the first comparison of P.C. VI, in 1886, they would have reached an effectively steady condition, but they would, in the meantime, have shortened by just over 0·0002 in. since their own original determinations in 1852.

With regard to the metre, the present international standard, when given legal status in 1889, was declared equal, within the limits of accuracy of measurement estimated at the order of 0·001 mm., with the old Mètre des Archives which it superseded. Evidence of subsequent consistency is available from the results of recomparisons of twenty national copies since 1919 with the working standards of the Bureau International.† After making due allowance for changes known to have occurred in the values of the working standards, and for certain adjustments in the accepted values of the coefficients of expansion of the various bars, the results of all these individual comparisons were in agreement within ± 0·001 mm. with the original determinations made prior to 1889, the average change being ± 0·0003 mm., and the mean of all the results was in agreement with the original mean within 0·0001 mm. As regards the secular constancy of the metre, there have been a series of determinations of its length in terms of the wave-length of the red radiation of cadmium, commencing in 1893 with that of Michelson & Benoît (1895), which, while they exhibit variations amounting to ± 0·0003 mm. from the mean, give no indication of any progressive change. There is therefore good reason for confidence in the constancy of this unit.

---

* 1936, *Report by the Board of Trade on the comparison of the Parliamentary Copies of the Imperial Standards.*

† 1927, *La Création du Bureau International et son Œuvre*, p. 125, Ch. Ed. Guillaume. Paris: Gauthier-Villars.

The current accuracy of measurement of the yard is estimated at about 0·00002 in., or 0·5 part in 10^6, that of the metre, which has more perfect graduations, at about half this amount. Recently a number of the copies of the metre at the Bureau International have been regraduated, with still better lines, and an accuracy of the order of 1 part in 10^7 is now claimed. Fortunately, the prototype metre itself has particularly good lines.

Types of standard

It will be noticed that the standard yard, originally an end-bar, passed through the phases of the 'dot' standard, and the 'natural' standard defined by reference to the 'seconds' pendulum, and finished as a line-standard. The history of the metre is remarkably similar, though the order is somewhat different. Originally a 'natural' standard, defined as the ten-millionth part of the earth's polar quadrant passing through Paris, and represented for practical purposes by the Mètre des Archives, a platinum end-standard 25 × 4 mm. in section, it was fairly soon found that the difficulty of repeating the determination of the value of the standard by reference to the definition was so great that it was necessary to legalize the length of the Mètre des Archives itself as defining the unit. Later, when the new international standard of the metre was established, the principle of defining the unit by means of a material standard was maintained, but preference was given to a line-standard, graduated on the neutral plane, as had already been adopted by British scientists for the yard.

The parallelism between the histories of the British and metric standards of length is worthy of some comment. It is evident that British scientists have not been behind their Continental colleagues in the initiation of improvements. The common change from end- to line-standards, and the common abandonment of natural standards, is clearly indicative of the stage of technical development reached at the epochs concerned. When the earliest standards were constructed end-standards offered greater advantage in accuracy than line-standards. Later, with the application of the ruling engine and microscope, this position was reversed. But the story is not yet complete. There are now very good prospects of being able in the near future to propound a natural standard, in the form of a selected wave-length of monochromatic light, which is capable of definition and reproduction with an accuracy of the order of 1 or 2 parts in 10^8—appreciably higher than can be attained even with the best line-standards. Further, there are now technical means available to produce end-standards (which are those most readily and directly measurable in terms of wave-lengths) of sufficient perfection to take advantage of this accuracy.

It may be of interest to remark here that the pendulum experiment, regarded as a means of ascertaining the value of gravity in terms of the units of length and time, still remains one of the essential operations of metrology, since a knowledge of the local value of gravity is necessary, for example, to the establishment of the electrical unit of the ampere by weighings on a current balance, and of the standard atmospheric pressure in determining the boiling-point of water for fixing the temperature scale. Mr Clark's note in the present series of papers summarizes the position with regard to the accuracy with which the value of g has been determined.

Ratio of yard to metre

It might be supposed that further evidence as to the constancy of the yard would be available from successive determinations of its relationship to the metre. Unfortunately, however, early determinations of this ratio, by Kater (1818) and Clarke (1867 & 1873) were made with standards about which adequate knowledge is now lacking, so they are not very reliable for this purpose. Three determinations have, however, been made of the ratio of the present Imperial Standard Yard and International Prototype Metre, the results being as follows:

1 m. = 39·370113 in. (Benoît & Chaney 1895),
1 m. = 39·370147 in. (Sears, Johnson & Jolly 1928),
1 m. = 39·370138 in. (Sears & Barrell 1934),

the last of these being an indirect value obtained from independent determinations of the yard and metre in terms of the wave-length of light. With regard to these it can be said that although they suggest a possibility that the yard has shortened slightly in relation to the metre during the past 40 years, the whole range of the variations exhibited does not exceed the possible range of experimental error in the comparisons, so that the previous conclusion as to the present stability of the yard may still be regarded as justified.

While on the subject of the relationship between the yard and the metre it may be of interest to mention that, by an Act of Congress of 1866, legal sanction was given in America for the use of the relationship

1 m. = 39·370000 in.,

and that by a subsequent order of the United States Treasury, known as the Mendenhall Order, this ratio is now accepted as defining the American inch in terms of the International Metre, instead of vice versa. It follows that there is at present a small difference between the U.S. and U.K. inches, the former being equivalent to 1 in. = 25·400051 mm., and the latter (taking the mean of the three determinations above mentioned) to 1 in. = 25·399965 mm. These two values are both so near to the convenient round value 1 in. = 25·4 mm. that it would be very advantageous if adjustment could be made in both countries to agree on this common ratio for the future by redefining the yard as equal to 0·9144 m. exactly.

It may be remarked that if the conclusion suggested above as to the shortening of the Imperial Standard Yard be correct, then the ratio of the Imperial Yard to the International Metre (had the latter been already in existence at the dates mentioned) would have passed through the value 1 m. = 39·370000 in. about the year 1858 (very soon after the legalization of the British standard) and, curiously enough, through the value 1 m. = 39·370079 in., corresponding to 1 in. = 25·400000 mm., about 1865, just before the former value was adopted in America.

Until comparatively recently the difference has been of no practical significance, but it is now just appreciable in relation to the accuracy guaranteed by the makers of the most accurate end-gauges used in industry. The round figure has the further practical advantage of facilitating mechanical conversion by means, for example, of a 127-tooth gear wheel on a lathe or measuring instrument, between the units of the two systems.

Comparison of line- and end-standards

Whether the primary material representation of the unit of length takes the form of a line-standard or of an end-standard, the operation of transferring from the one to the other is an inevitable necessity, since both types are needed for practical applications in different fields of measurement. Probably the best method of carrying out this comparison which has so far been devised is the following.

Taking, for example, the case of the yard, it is necessary to provide an auxiliary end-bar of say 35 in. in length, with ends highly finished and accurately plane and parallel. In addition, two small end-blocks each 1 in. in length are required, similarly finished as regards their ends, with graduation marks on plugs inserted in holes at the level of their neutral planes, similar to those on the standard yard. The perfection of the surfaces will enable these end-blocks to be adhered to the ends of the 35 in. bar by the process known as 'wringing'. If both the end-blocks are so attached the combination constitutes a built-up line bar which can be compared with a 36 in. line-standard in the comparator in the usual manner. If only one of the end-blocks at a time is attached to the 35 in. bar the combination can be compared with a 36 in. end-standard on an end-measuring machine. Let the length of the 35 in. bar be B, that of the line-standard L, and that of the end-standard E. And let the distances of the graduation lines from the surfaces of the two end-blocks be a_1, b_1, a_2, b_2 respectively. Finally, let the thickness of a 'wringing film' be t. If the observed differences are x_1, x_2 and y_1, y_2, the results of the measurements may then be written as follows:

$$B+a_1+a_2+2t = L+x_1, \quad B+b_1+b_2+2t = L+x_2,$$

whence
$$B+\tfrac{1}{2}(a_1+b_1+a_2+b_2)+2t = L+\tfrac{1}{2}(x_1+x_2),$$

and
$$B+a_1+b_1+t = E+y_1, \quad B+a_2+b_2+t = E+y_2,$$

whence $$B+\tfrac{1}{2}(a_1+b_1+a_2+b_2)+t=E+\tfrac{1}{2}(y_1+y_2).$$

Thus, finally, $$E+t=L+\tfrac{1}{2}(x_1+x_2)-\tfrac{1}{2}(y_1+y_2),$$

whence it will be seen that the lengths of both the 35 in. bar and of the end-blocks have been eliminated.

Actually, of course, a number of variations can be made in the wringings which are not included above. In practice all possible variations would be employed, and the mean result taken. But the principle of the comparison should be clear. The fact that the thickness of one wringing film is associated with the length of the end-standard as the result of this comparison corresponds to the fact that, in practical use, precision end-gauges are habitually combined together, by the process of wringing, so as to make up composite standards of any desired size. The calibrated mechanical size of each gauge must therefore always include the thickness of one wringing film, which may be regarded as distributed half over each end of the gauge, so that when wrung together in groups the thickness of one wringing film is automatically provided for at each interface. In addition, the composite gauge, just as a single gauge representing the same size, will be associated with one film thickness divided between its terminal faces, and allowance may need to be made for this in some circumstances, e.g. when deriving the 'mechanical' length of an end-standard from its 'optical' length under certain conditions. The actual thickness of an individual wringing film is, however, very small, amounting only to about 0·0000002 in. (Rolt & Barrell 1927).

Determination of submultiple end-standards

Now follows the process of determining the values of end-standards representing subdivisions of the fundamental unit, which are of great importance in practical applications. This can be done by comparing together, in a suitable end-measuring machine, combinations of gauges of equal nominal size. It is convenient, though not essential, that a duplicate series should be available. As a typical case consider the entries in table 2, which refers to the determination of the values of a series of end-gauges 7, 8, 9, 10 and 11 in., in relation to those of 6 and 12 in. supposed already known. Duplicate gauges are indicated by a dash (′).

TABLE 2

12+10	11+11′	...	...	...	...
12+ 9	11+10	10+11′	...	...	...
12+ 8	11+ 9	10+10′	9+11′	...	...
12+ 7	11+ 8	10+ 9	9+10′	8+11′	...
12+ 6′	11+ 7	10+ 8	9+ 9′	8+10′	7+11′
6+11′	11+ 6′	10+ 7	9+ 8	8+ 9′	7+10′
6+10′	...	10+ 6′	9+ 7	8+ 8′	7+ 9′
6+ 9′	...	...	9+ 6′	8+ 7	7+ 8′
6+ 8′	...	...	...	8+ 6′	7+ 7′
6+ 7′	...	...	...	...	7+ 6′

It will be noted that all the combinations in any row of this table have the same nominal length. Suppose each combination in any column is compared with the corresponding combination in column 1, and let the observed differences be x_1, x_2, etc. Adding all the results relating to any particular column it is seen that the sum of the right-hand members in the combinations compared cancels out, and there remains a series of equations of the form, taking, for example, column 3,

$$10\times 6 = 12\times 4 + 6\times 2 + \Sigma x \quad \text{or} \quad 10 = 12\times 4/6 + 6\times 2/6 + \Sigma x\times 1/6.$$

This case has been selected as exhibiting in a simple form the general procedure involved. It is easy to follow the modifications involved, say, in deriving a series such as 30, 24, 18, 12 and 6 in. from a yard end-standard, or a series such as 0·100, 0·101, 0·102, ..., 0·110 in. which is found in a normal set of slip gauges. Reference gauges of the smaller sizes are, however, more readily measured, provided the flatness and parallelism of their surfaces is sufficiently high, by interference methods in terms of the wave-length of light, other gauges being then compared directly with reference gauges of similar size in a suitable end-measuring machine.

By one or other of the above processes the lengths of reference end-gauges can be established to an accuracy of 1 part in 10^6 for sizes above 1 in., or to 0·000001 in. for sizes below. It may be worth while to point out that this accuracy is actually necessary to meet industrial demands at the present time. Grouped combinations which may include as many as five such gauges are required to be correct to the nearest 0·0001 in., that is, to within 0·00005 in., and their users expect to be able to rely on this without making use of corrections for known errors. Each individual gauge must therefore be correct within 0·00001 in., and the makers guarantee this, and in some cases even higher, accuracy. Reference gauges used in verifying that such guarantees have been fulfilled must therefore be known at least to 0·000001 in. It follows that the present difference of nearly 4 parts in 10^6 between the British and American inches is, as already stated, no longer negligible in relation to the highest grade of current commercial accuracy.

Subdivision of line-standards

To complete the present survey reference must be made to the operations of calibrating the subdivisions of a divided scale, and of building up from the yard or metre to the longer standards which are required, for example, in field-survey work.

For the first of these processes the bar to be examined is mounted, under two micrometer microscopes, on the girder of a comparator in which the carriage, instead of moving transversely, is arranged to move in the direction of the length of the bar. The separation of the microscopes can be varied to suit the successive stages of the comparisons. Suppose, for example, one wishes to determine the decimetre subdivisions of a 1 m. bar. The microscopes are set first 1 dm. apart and direct comparisons are made between the 0/1 interval and the 1/2, 2/3, etc., intervals up to 9/10. The results of these measurements give a series of values, $a_1, a_2, a_3, \ldots, a_9$ for the differences 0/1–1/2, 1/2–2/3, ..., 8/9–9/10 between the successive adjacent decimetre intervals on the bar. The microscopes are then placed 2 dm. apart, and similar comparisons made of the intervals 0/2, 1/3, etc., up to 8/10. The results of these measurements, subtracting the parts common to the two adjacent intervals compared, give a series of values $b_1, b_2, \ldots, b_8$ for the differences 0/1–2/3, 1/2–3/4, etc., up to 7/8–9/10 and so on, with successively increasing distances between the microscopes, up to 9 dm. The results of all these measurements can then be tabulated as follows:

	0/1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10
0/1	0	$-a_1$	$-b_1$	$-c_1$	$-d_1$	$-e_1$	$-f_1$	$-g_1$	$-h_1$	$-k$
1/2	a_1	0	$-a_2$	$-b_2$	$-c_2$	$-d_2$	$-e_2$	$-f_2$	$-g_2$	$-h_2$
2/3	b_1	a_2	0	$-a_3$	$-b_3$	$-c_3$	$-d_3$	$-e_3$	$-f_3$	$-g_3$
3/4	c_1	b_2	a_3	0	$-a_4$	$-b_4$	$-c_4$	$-d_4$	$-e_4$	$-f_4$
4/5	d_1	c_2	b_3	a_4	0	$-a_5$	$-b_5$	$-c_5$	$-d_5$	$-e_5$
5/6	e_1	d_2	c_3	b_4	a_5	0	$-a_6$	$-b_6$	$-c_6$	$-d_6$
6/7	f_1	e_2	d_3	c_4	b_5	a_6	0	$-a_7$	$-b_7$	$-c_7$
7/8	g_1	f_2	e_3	d_4	c_5	b_6	a_7	0	$-a_8$	$-b_8$
8/9	h_1	g_2	f_3	e_4	d_5	c_6	b_7	a_8	0	$-a_9$
9/10	k	h_2	g_3	f_4	e_5	d_6	c_7	b_8	a_9	0

It will be noted that in this table each entry is repeated, above the left to right diagonal, with reversed sign. Taking any particular column of the table, for instance the fourth, and adding up all the entries, it is seen that the total of all the observations relevant to the interval 3/4 leads to the result

$$10 \times 3/4 - \Sigma(0/1 + 1/2 + \ldots + 9/10) = -(c_1 + b_2 + a_3) + (a_4 + b_4 + \ldots + f_4),$$

or

$$3/4 = \tfrac{1}{10} \times 0/10 - \tfrac{1}{10}(c_1 + b_2 + a_3) + \tfrac{1}{10}(a_4 + b_4 + \ldots + f_4),$$

and similarly for any other interval.

With slight modification the same process can be applied to determining the values of every centimetre of two particular decimetres in terms of the whole length of the other. In this case each space in the square will be occupied by an independent observational result. The millimetres of two particular centimetres can then be determined in the same way.

The accuracy of this process, given good conditions, is remarkably high, since the bar may be assumed to have the same coefficient of thermal expansion throughout its length, and errors due to temperature variations, if kept within reasonable limits, should be negligible. With good lines it is considered that a final accuracy of the order of 0·0002 mm. in the placing of any intermediate subdivision in relation to the terminal graduations can be readily attained.

Longer standards

Turning to the converse process, of building up to larger lengths, the first step is the comparison of a long bar, say of 4 m. length, 1 m. at a time with a reference standard metre, in a transverse comparator of suitable capacity. As this bar will subsequently be used in air, under less favourable conditions for temperature control, it is nowadays usually made of invar. When standardized it is used either as a basis of comparison for 4 m. intervals on a mural base or, alternatively, the bar may be directly compared with 4 m. intervals graduated on the surface of a standard invar tape. These normally have a total length of 24 m.

Tapes intended for use in the field in setting out bases for geodetic surveys are compared either with the appropriate interval on the mural base, or directly with a standard tape of similar length. In the case of English measures the most usual length is 100 ft., built up in eight steps from a bar 12 ft. 6 in. in length. The operation of standardizing this bar in terms of the yard consists (*a*) of determining each 6 in. length in terms of the whole length by the method of subdivision, and (*b*) determining the lengths of the 0 to 12 ft. and 6 in. to 12 ft. 6 in. lengths in four steps against a reference standard yard.

It is considered that under the best conditions an accuracy of the order of 1 part in 10^6 can be attained in the standardization of the field tapes in either the British or the metric systems. Field surveyors find that, even in the adverse conditions of open-air working, they are often able to repeat their measurements of a long base to appreciably higher accuracy than this, and are sometimes inclined to criticize the accuracy offered them as the result of calibration under laboratory conditions. The explanation is, of course, that so many observations are taken in the field that casual errors are averaged out, and the probable error of determination of the whole length of, say, a 20 km. base in terms of the length of a particular tape, becomes very small. In the laboratory, on the other hand, we are concerned, not merely with the repeatability of observations, but with their relationship to the absolute value of the unit. Some countries have already experimented with the possibility of determining the length of a mural base in terms of the wave-length of light, and it is intended to do so also at the National Physical Laboratory when opportunity affords. So far, however, we have no experience of the method in this special application.

It may be of interest, finally, to record that the 50 m. mural base at the National Physical Laboratory has gradually grown, since it was first built in 1908, by something over a centimetre, due presumably to progressive chemical action in the cement with which the bricks are bonded.

THE STANDARDS OF LENGTH IN WAVE-LENGTHS OF LIGHT

BY H. BARRELL

Historical survey

The first direct measurement of the metre in terms of wave-lengths of the red radiation of cadmium was made by Michelson & Benoît (1895) in 1892–3 at the Bureau International des Poids et Mesures (B.I.P.M.). In 1905–6, Benoît, Fabry & Perot (1913) repeated the determination of the metre, using improved methods and apparatus. The International Solar Union (now the International Astronomical

Union) adopted, in 1907,* the wave-length of the red line of cadmium as the reference basis for all spectroscopic measurements of wave-lengths of light, and defined its value in terms of the International Ångström. Using the relationship between the wave-length of the red line of cadmium originally quoted by Benoît *et al.* namely, $\lambda_R = 6438{\cdot}4696 \times 10^{-10}$ m. in normal† air, the Ångström was defined by assigning the value of 6438·4696 A to the wave-length of the red line of cadmium in normal air, and hence can differ only to the extent of the error in the above determination from 1×10^{-10} m.

In 1923, the International Committee of Weights and Measures‡ accepted in principle the possibility of the eventual adoption of a definition of the metre in terms of wave-lengths of light, subject to the formulation of satisfactory conditions for practical realization of such a definition, and urged national standards laboratories to undertake investigations for this purpose. Four years later, the International Conference on Weights and Measures§ (1927) gave formal sanction, as an interim measure, to determinations of length being made by reference to the wave-length of the cadmium red radiation, determined by Benoît *et al.*, as an alternative to direct reference to material standards.

Reports of the Board of Trade (1930, 1936)‖ have expressed the need for replacing the present Imperial Standard Yard by a standard more in accordance with modern requirements, and also suggest the possibility of defining the yard in terms of wave-lengths.

Following the International Committee's recommendation of 1923, the first result to be announced was that of Watanabe & Imaizumi (1928); it was the outcome of work carried out by means of apparatus, constructed in this country for the Japanese Government, which was almost identical with that of Benoît *et al.*

Each determination so far mentioned made use of purely line-standard technique by relating the position of a line or lines engraved on an interferometer standard (or étalon) to the graduations on a copy of the metre by microscope observations. With the introduction of modern types of end-standard constructed in the form of steel blocks or bars in lengths up to a metre or more and having flat, parallel, terminal faces with a mirror-like finish of optical perfection, it became possible to apply interferometric methods, as devised by Michelson and others, directly to the measurement of a practical standard of length.

Determinations of the metre and the yard made by Sears & Barrell (1932, 1934) at the National Physical Laboratory (N.P.L.) and of the metre by Kösters &

---

* 1907, *Trans. int. Un. Sol. Res.* (2), **20**, 28.

† The modern specification of normal air for spectroscopy and metrology is: dry air at 15° C under a pressure of 760 mm. of mercury at 0° C ($g = 980{\cdot}665$ cm./sec./sec.) and containing a normal proportion of 0·03 % by volume of carbon dioxide usually found in the open air and in well-ventilated rooms.

‡ 1923, *P.V. Com. int. Poids Mes.* p. 67.

§ 1927, *P.V. Com. int. Poids Mes.* p. 67.

‖ 1930 and 1936, *Report of the Board of Trade on the comparisons of the Parliamentary Copies of the Imperial Standards.* H.M.S.O.

Lampe* at the Physikalisch-Technische Reichsanstalt (P.T.R.) in 1933 produced evidence of this trend, for the interferometry at both institutions was concerned with the measurement of an end-gauge approximately equal to one of the fundamental units. Correlation of the optically measured end-gauges with the fundamental line standards was effected by the already well-established metrological procedure for transferring from line- to end-measures,† to which reference is made in Mr Sears's paper in the present series.

Other important changes were introduced in both the new methods. With the N.P.L. apparatus it was possible to measure the end-gauge either in air under known and controlled conditions or *in vacuo*; with the P.T.R. apparatus the end-gauge was measured in air and a concurrent determination made of refractive index over an optical path equivalent to that of the gauge length. Thus the results obtained from both methods can be expressed, if desired, in terms of wave-lengths *in vacuo*, rather than in air for which the temperature, pressure and composition must also be specified in order to define the refractive index. Platinum resistance thermometry was used in both determinations for the measurement of temperature of the end-gauge. In the neighbourhood of 20° C it is probable that the temperature scale defined by a platinum resistance thermometer conforming with international specification does not differ by much more than 0·001° C from the thermodynamic Centigrade scale, corresponding to an uncertainty of only about $0{\cdot}01\mu$ in the length of a 1 m. end-gauge constructed of steel.

Owing to the disappointingly large discrepancy between the originally announced N.P.L. and P.T.R. results, the determinations were repeated in 1934–5 (Kösters & Sears 1935), but the mean discrepancy remained equivalent to $0{\cdot}54\mu$ in the metre. At the same time, however, direct optical comparisons of the N.P.L. and P.T.R. metre end-standards, both independently established in terms of light waves *in vacuo*, were made at both institutions, and the results of all the optical measurements were found to be mutually consistent to ± 1 part in 36 millions, equivalent to $0{\cdot}028\mu$ in the metre.

Results of subsequent P.T.R. work have recently been noticed (Kösters 1938) which indicate that the former P.T.R. results have been modified in a sense which considerably reduces the original divergence, and this change is substantiated by a new determination made in 1937, the result of which differs from the more recent of the two N.P.L. results by an amount equivalent to only $0{\cdot}14\mu$ in the metre.

In 1940, a preliminary determination of the metre was made in Russia (Barinov 1941) by methods which appear to be similar to those used in Germany.

Values of the yard have been determined by Tutton (1931) and from the N.P.L. work mentioned above.

It should be mentioned that Williams (1933) has proposed a novel method of measuring an end-gauge by comparison with a reflexion echelon grating constructed of fused silica and of length approximately equal to the metre or the yard.

---

* 1934, *C.R. Conf. Gén. Poids Mes.* p. 79. Paris: Gauthier-Villars.

† 1919, *Ann. Rep. Nat. Phys. Lab. Lond.* p. 90.

Wave-length measurements of length made under generally practicable conditions in air effectively involve, for their reduction to comparable terms, concurrent observations of refractive index of the air. Such observations may be made either directly, as in the P.T.R. apparatus, or indirectly from a knowledge of the temperature, barometric pressure, humidity and possibly the CO_2 content of the air and of how these variable factors influence the refractive index. For spectroscopic purposes use is made of the experimental data obtained in 1918 by Meggers & Peters (1918–19) on the refraction and dispersion of air. At the 1931 meeting of the International Committee of Weights and Measures it was decided that further precise measurements of the refractive index of air were required. Concordant results for the refraction and dispersion of air in the visible region have since been announced from the B.I.P.M. (Pérard 1934), the P.T.R. (Kösters & Lampe 1934) and the N.P.L. (Barrell & Sears 1939), which appreciably disagree with the hitherto accepted tables of Meggers & Peters for the same spectral region. Suggestions have been made (Sears & Barrell 1934; Kösters 1938) for defining the reference standard of wave-length in the vacuum condition and so eliminating the refractive index entirely from the fundamental definition.

With one exception, all wave-length determinations of the metre have been made directly in terms of the red radiation of cadmium, which was selected by Michelson as the most suitable for the purpose. Certain radiations of krypton have since been proposed as better alternatives, notably the yellow-green line (λ5651 A) by Kösters* which was used during the actual P.T.R. determinations of the metre. The P.T.R. results were finally expressed, for comparison with other work, in terms of the cadmium red radiation by use of an experimentally determined relation between the two wave-lengths.

Discussion of results obtained on the metre and the yard

(*a*) *The metre*. The original results of all the determinations mentioned in the previous section are not strictly comparable, as the experimental conditions were not identical, nor in all cases fully specified. But, reducing them as far as possible from the data available to a common basis, the following table has been prepared.

The values originally quoted by the observers are corrected in the fourth column of the table below to take account of subsequent conclusions regarding the values to be attributed to the standards of length employed and adjusted, so far as the information available permits, to uniform conditions of normal air (see footnote, p. 165). It is to be noted that no evidence of any systematic change in the length of the metre is shown, over a period of nearly 50 years, compared with the wave-length of light; also the greatest departure of any individual result from the mean is less than 0·3 in 10^6, which is of the same order of accuracy as that generally associated by experienced metrologists (Sears 1936; Kösters 1938) with the measurement of length in terms of the existing national copies of the metre.

---

[*] 1927, *P.V. Com. int. Poids Mes.* p. 67.

Furthermore, the mean value of the wave-length is almost identical with that originally quoted by Benoît *et al.* which has been accepted as the basis of spectroscopic measurements of wave-lengths and involves the definition of the Ångström.

The mean of the two N.P.L. values of the wave-length of the red line *in vacuo*, determined respectively in 1933 and 1934–5, was $6440{\cdot}2513 \times 10^{-10}$ m.‡ (Kösters & Sears 1935). If the mean corrected and adjusted P.T.R. value of $6438{\cdot}4693 \times 10^{-10}$ m. for normal air, abstracted from the table, is reduced to vacuum condition by use of the P.T.R. value of 1·000,276,47 (Kösters & Sears 1935) for the refractive index of the red line in normal air, the mean P.T.R. value of the wave-length *in vacuo* becomes $6440{\cdot}2493 \times 10^{-10}$ m.

Values of the wave-length of the cadmium red radiation in terms of the International Metre (*unit* $= 1 \times 10^{-10}$ *m.*)

Date of Determination	Observers	Original Values	Corrected and adjusted Values in normal air	Differences from Mean	
				10^{-10} m.	parts per 10^6
1892–3	Michelson & Benoît	6438·4722	6438·4691*	−0·0004	−0·06
1905–6	Benoît, Fabry & Perot	6438·4696	6438·4703*	+0·0008	+0·12
1927	Watanabe & Imaizumi	6438·4685	6438·4682	−0·0013	−0·20
1933	Sears & Barrell	6438·4711	6438·4713	+0·0018	+0·28
1934–5	Sears & Barrell	6438·4709	6438·4709	+0·0014	+0·22
1933	Kösters & Lampe	6438·4672	6438·4689	−0·0006	−0·09
1934–5	Kösters & Lampe	6438·4685	6438·4690	−0·0005	−0·08
1937	Kösters & Lampe	6438·4700	6438·4700	+0·0005	+0·08
1940	Romanova, Varlich, Kartashev & Bartarchukova†	6438·4677	6438·4677	−0·0018	−0·28
		mean	6438·4695	±0·0010	±0·16

* See Guillaume (1927). † See Barinov (1941).

(*b*) *The yard.* The number of wave-lengths of the cadmium red radiation contained in the yard under normal air conditions, as determined by Tutton (1931), was 1,420,209·8.

The following N.P.L. results§ (Sears & Barrell 1934) were obtained in 1933, using the red radiation of cadmium:

$$1 \text{ yard} = 1{,}419{,}818{\cdot}24 \text{ wave-lengths } \textit{in vacuo}$$
$$= 1{,}420{,}210{\cdot}74 \text{ wave-lengths in normal air.}$$

---

‡ The corresponding number of wave-lengths in the metre is 1,552,734·44, which may be associated with the value of the yard in terms of wave-lengths *in vacuo* to determine the relationship between the yard and the metre, as quoted later.

§ These values are slightly lower than those originally published owing to a reduction of 0·07 wave-length in the value previously assigned to the correction applied to the optically measured end-gauge, in the N.P.L. method, to derive the practical mechanical length (Rolt & Barrell 1929). This small change has been made as the result of further work on determining the phase loss in light reflected at the lapped steel surfaces of the gauge.

The difference between the Tutton and N.P.L. values in normal air accords with the precision of measurement actually attainable with the present Imperial Standard Yard, which is assessed at about 0·5 part in 10^6 (Sears 1936).

The relationship of the yard to the metre, derived from the comparison of the N.P.L. wave-length determinations quoted above, is given by

$$1\text{ m.} = 39{\cdot}370138\text{ in. (or 1 in.} = 25{\cdot}399962\text{ mm.).}$$

Future definition of the units of length

From the foregoing, it is clear that it may soon be possible to supersede the present definitions of the fundamental units of length, based on the distances between lines engraved on certain metal bars, by a definition in terms of the natural standard provided by the wave-length of a monochromatic ray of light. The various national laboratories already employ interferometric methods in routine procedures (Kösters 1938; Barrell 1945) for directly calibrating the practical reference standards of measurement—such as, for instance, sets of slip or block gauges containing sizes up to 10 cm. or 4 in.—in terms of this natural, invariable and easily reproducible unit of length to an accuracy of $0{\cdot}025\mu$ or one-millionth of an inch. Experience in this use of a wave-length standard shows that it is capable of a higher precision and greater convenience in the measurement of such practical standards than the line-standard. The results of the British and German experiments on the calibration of end-gauges 1 m. in length demonstrate that it is possible to establish metre lengths independently in two countries on the basis of a wave-length standard of reference with a precision of $\pm 0{\cdot}03\mu$, which is approximately ten times better than is generally attainable with the present material standards. If the definitions of the units of length were established in terms of wave-lengths the end-standard would obviously become the first material representation of both units, rather than the line-standard.

The red radiation of cadmium, selected 50 years ago by Michelson as the best available, is still the approved standard for all measurements of wave-lengths of light, but it is not now considered to be quite as ideally monochromatic as is desirable for a universal standard of length. Cadmium is known to consist of a mixture of six isotopes, four being of even and two of odd atomic masses; krypton, the suggested alternative source of monochromatic radiation, also consists of six isotopes, five being of even atomic masses and one odd. It is found that the hyperfine structure of certain radiations emitted by cadmium and krypton may be accounted for by associating a nuclear spin of $I = \frac{1}{2}$ with each of the odd isotopes of cadmium and of $I \geqslant \frac{7}{2}$ with the odd isotope of krypton. Some evidence of hyperfine structure, due to the odd isotopes, has, indeed, been observed in the red line of cadmium, but none has so far been observed in the yellow-green line of krypton, possibly because of its rather low intensity. Although the even isotopes of both elements, by reason of their zero nuclear spin, are not expected to produce complexity, there may be some minute mass displacement effect between the corresponding radiations emitted by the different even isotopes in each element.

The conditions under which excitation of an element takes place in the discharge lamp are also known to affect the monochromatic quality of the emitted radiations, and care must be taken to reduce, as far as possible, perturbative effects due to Doppler motions of the radiating atoms, electric and magnetic fields (Stark and Zeeman effects), collisions, pressure and reversal. As regards the Doppler effect of motions in a radiating gas or vapour, it is known, from kinetic theory, that the 'width' of a line is proportional to $\sqrt{(T/A)}$, where T is the absolute temperature and A the atomic mass, so that radiation tends to become more monochromatic when excited from heavy atoms at low temperature.

A new possibility now exists of finding a more suitable radiation from a single isotope of a heavy element with zero nuclear spin (and therefore with even mass number), owing to the recent advances in the technique of separating isotopes. If a single even-number isotope of mercury, for instance, could be isolated and caused to emit the well-known green radiation 5461 A, either in a Paschen and Schüler hollow-cathode type of discharge lamp at the temperature of liquid air or by very high-frequency excitation of a fine atomic beam of such mercury viewed normally to the direction of motion in the beam, it might provide an ideal source of monochromatic radiation. Information was received from America in 1941 that a suitable isotope of mercury had been obtained from a radioactive isotope of gold.

Although there may be some minor objections in practice to adopting a wave-length *in vacuo* as the basis of definition, it is considered that the weight of ultimate scientific advantage is greatly in its favour. Whatever radiation may be eventually selected for this purpose, the new definition of the metre should preferably be so adjusted as to yield a value of $6438{\cdot}4696 \times 10^{-10}$ m. for the wave-length of cadmium red radiation in normal air, which happens to be almost identical with the mean experimentally determined value (see table), thus preserving the definition of the Ångström which from thenceforth could be redefined simply as 10^{-10} m. The yard could then with advantage be defined in similar terms to the metre, on a basis which would make the yard exactly equal to

$$0{\cdot}9144 \text{ m. (or 1 in.} = 25{\cdot}4 \text{ mm.).}$$

With the improvements that are now believed to be possible in producing pure monochromatic radiation, combined with new definitions based on wave-lengths *in vacuo*, it is considered that the units of length could be reproduced through the agency of end-standards with a precision approaching 1 part in 10^8. The best precision that could be expected from the existing definitions in terms of material line-standards, utilizing all modern resources for their production and observation, is unlikely to be superior to 1 part in 10^7.

THE STANDARDS OF MASS

BY F. A. GOULD

In this review it is proposed to consider briefly the various standards of mass, including primary, secondary and also the more ordinary standards in everyday use in trade, industry and scientific work. Particular reference will be made to the principal standards, their comparisons, and the technique of this class of weighing. Consideration will also be given to air buoyancy in relation to the derivation of secondary and other standards from the primary standards, and to conventions for the adjustment of ordinary standards.

Primary standards

The fundamental standard in the British system of units is the Imperial Standard Pound (Weights and Measures Act, 1878). This was constructed in 1844 in the form of a cylindrical piece of platinum of diameter slightly smaller than its height, and with a shallow circular groove around its flank to take an ivory lifting fork. There are five Parliamentary Copies of this standard, four of which are of platinum and one of platinum-iridium alloy, which serve for the replacement of the Imperial Standard in the event of its being lost or damaged. It is to be noted that this standard is the avoirdupois pound containing 7000 gr. The troy pound, which was the sole standard from 1824 to 1834, is no longer legal and the troy ounce, defined as 480 gr. is now the basis of the troy system.

In the metric system it is of interest to refer to the original conception underlying the definition of the unit of mass. The kilogram was originally defined by reference to a 'natural' standard, i.e. the mass of the cubic decimetre of water. The material representation of this standard was the Kilogramme des Archives, a simple cylindrical piece of platinum which was constructed in the latter part of the eighteenth century. The practical realization of the definition of the kilogram proved to be insufficiently precise, and it was accordingly superseded by the material standard itself, viz. the Kilogramme des Archives. Though there may be natural standards of mass which might be employed, there is no prospect as yet that any of these could be realized with a precision approaching that with which material standards of mass can be compared.

The Kilogramme des Archives remained the standard until 1889, when, under the auspices of the Metric Convention, it was replaced by the International Prototype Kilogram, the present fundamental standard in the metric system. The latter takes the form of a simple cylinder, of height equal to its diameter, and is made of the same 10 % platinum-iridium alloy as the international metre. At the time of this replacement the new kilogram was declared* to be equal to the earlier standard, to within the limits imposed by the uncertainty of the volume of the older standard, which had not been established so closely as that of the later one.

---

* 1889, *C.R. Prem. Conf. Gén. des Poids et Mes.* Paris; Gauthier-Villars.

Some forty or more copies of the international kilogram were constructed. The principal copies, or 'témoins', and other working copies are preserved at the International Bureau of Weights and Measures, Sèvres, Paris. Other copies are distributed, as national reference copies, to the various countries signatory to the Metric Convention. One such copy, No. 18, designated the British Copy of the Kilogram, is kept in this country.

Balances for comparing standards

Consideration will next be given to the balances used for the comparison of precision standards. The accuracy attained with an ordinary type of knife-edge balance of good quality is surprisingly high, and a precision of the order 1 part in 10^7 has for many years been attained in the comparisons of pounds or kilograms. The principle of double weighing is normally employed, each standard being weighed from each arm of the balance in turn. Due allowance is made for the atmospheric buoyancies of the standards, which are computed from a knowledge of the volumes of the standards, as experimentally determined, and the air density determined at the time of each weighing. The densities of the standards are determined by weighing them in air and in water. When a new standard is constructed, its density may with advantage be determined shortly before the final polishing of the standard.

Balances of special design for this class of weighing were in use on the Continent 50 years ago. These were provided with arrangements for interchanging the standards undergoing comparison without opening the shutters of the balance, thus enabling a double weighing to be made with the minimum of disturbance to the temperature of the balance. Balances of this type, constructed by the well-known firm of Rüprecht, Vienna, gained a high reputation and have long been in use at the International Bureau of Weights and Measures, Paris.* An accuracy of the order 1 part in 10^8, or somewhat finer, was attained in the comparison of kilograms as the result of repeated weighings. In this country, however, no substantial progress was made until recent years, when a new balance, designed and constructed at the National Physical Laboratory, was installed. Some of the principal features of this balance will be described briefly here, including also a special technique of weighing which has enabled a striking improvement in accuracy to be made. Figure 1 (plate 6) gives a general view of the instrument with its cover removed. Provision is made for weighing masses up to 1 kg. in air or *in vacuo*, though there is no intention of carrying out *in vacuo* any standardizations which should properly be made under ordinary conditions of atmospheric pressure. The balance is installed in a closed vault and is so designed that the whole operation of weighing, including the interchange of the standards during a double weighing, can be controlled from outside the vault. To minimize wear on the standards two loose pans are provided, and once the standards are placed on these, they are not disturbed during the whole of a sequence of weighings. Beneath these pans each pan suspension has an aperture cut in it in the form of a cross, open towards the centre of the balance, and to interchange the pans on the suspensions a ring surrounding the centre of the base, and bearing two arms with suitably shaped ends which can pass through these apertures, is first raised, lifting the pans off their suspensions, then rotated through 180°, and again lowered. Allowance, of course, has to be made for the difference between the unloaded pans, which can be determined by direct weighing, as often as is found necessary. Alternatively, the standards can be weighed on each pan in turn, and the mean result taken.

---

* 1881, *Trav. Bur. int. Poids Mes.* **1**.

Another feature of the balance is the design of the two series of stirrups, supported on crossed knife-edges, through which the load is transmitted to the beam. These are arranged so that in combination they provide precise location of the loads transmitted by them on each terminal knife of the beam.

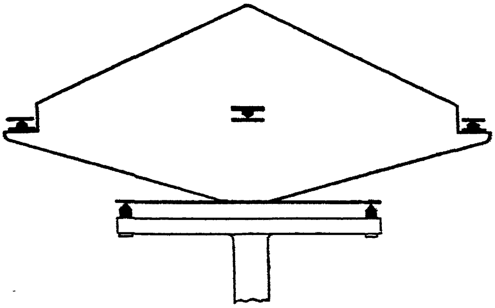


Figure 2. Diagram of arrangement for steadying the beam.

But even with these special features the improved accuracy which is available with the balance would not have been attained without the adoption of a special technique for carrying out the weighings. In ordinary circumstances a knife-edge balance would be fully arrested between successive determinations of its equilibrium point and also whenever the loads on the pans are being changed. During arrest the knives would be separated from the bearing planes. Many years ago Poynting (1878) showed that the residual errors associated with the use of the balance for fine weighings can be considerably diminished if the balance can be brought to rest, and the interchange of standards on the pans safely effected, while the knives are still in contact with their bearing planes. In the N.P.L. balance this principle is applied by directly steadying the beam after determining its equilibrium position. To enable this to be done, a slightly flexible metal strip is attached horizontally to the lowest part of the beam (figure 2). When the balance is swinging, a pair of agate points suitably mounted on an independent bridge below the beam can be raised so as just to meet the underface of the strip as the beam passes through the horizontal position, and so bring the balance temporarily to rest without separating the knives from their bearing planes. A separate control is provided for this operation, which can readily be carried out. In this condition of the balance, the pan-hooks can be held in arrest, and the pan plates interchanged without disturbing the beam and the stirrups. The balance is then set free to swing, and the double weighing completed without separating the knives and bearings. An accuracy closely approaching 1 part in 10^9 has been attained in the comparison of kilograms on this balance.

Stability of the pound and the kilogram

With regard to the stability of the primary standards of mass there is no 'natural' control to which reference can be made as in the use of the wave-length of light to check the standards of length. Deductions as to the stability of the pound and the kilogram have necessarily to be based on such evidence as may be found in the results of repeated comparisons of different standards over the period of time considered. So far as the pound is concerned, the results of intercomparison of the Imperial Standard Pound and its Parliamentary copies extend back to 1846, though only two intercomparisons were made prior to 1892. For various reasons the evidence is somewhat limited, but it clearly suggests that in relation to its copies the Imperial Standard Pound diminished by 1 part in 3½ million between 1846 and 1883 but remained constant to within about 1 part in 10 million between 1883 and 1933, when comparisons were last made.* The decrease in mass was attributed to wear resulting from too frequent usage. It should, however, be borne in mind that the Imperial Standard Pound and its original copies are somewhat inferior to the relatively modern kilogram standards in respect to the quality of the platinum and its surface finish. The Imperial Standard is the sole ultimate standard, whatever its mass may be; and only in the event of its replacement owing to loss or damage would the copies be used for reproduction of the standard.

Evidence as to the stability of the metric standard of mass will first be considered as from 1889, when the International Prototype Kilogram was installed. To avoid wear, this standard has been preserved unused for at least 50 years, and though arrangements have been made more recently for its recomparison with certain copies and also with the Kilogramme des Archives, the results are not yet available. Many of the copies have occasionally been re-verified by mutual intercomparisons, and the new values obtained for them, based on reference to certain standards at the Bureau (including the témoins), have been found to be in very close accord with those originally assigned to them (Guillaume 1927). Of sixteen copies so compared in 1889 and again in 1899, only four showed any change exceeding 0·02 mg. (viz. 2 parts in 10^8). Most of the copies which had been kept out of service between the two sets of intercomparisons showed only minute changes, comparable with the order of accuracy of weighing. Similar results were obtained for the interval 1899 to 1911, and so on. It is of interest to note the high stability of the British Copy

---

* 1936, *Report by the Board of Trade on the comparisons of the Parliamentary Copies of the Imperial Standards.*

(no. 18), which has been standardized at the Bureau on three occasions. The resulting values, which are given in table 3, do not differ among themselves by more than 2 parts in 10^8.

TABLE 3. BRITISH COPY OF THE KILOGRAM

Year	Mass (g.)
1889	1000·000 070
1924	1000·000 051
1933	1000·000 058

In addition to the evidence obtained from comparisons within the pound and kilogram systems respectively, there is also available,* as a check, a series of four well-established determinations of the relationship between the two systems, the results of which are given in table 4.

TABLE 4

Date	Value of the Imperial Standard Pound in grams	Referred to
1846	453·592 652	Kilogramme des Archives
1883	453·592 428	International Kilogram
1922	453·592 343	
1933	453·592 338	

These results, when plotted, lie on a smooth curve showing a continuous diminution of the pound, as compared with the kilogram. From 1846 to 1883, the diminution amounted to 1 part in 2 million, of which rather more than one-half was suggested by the evidence of the comparisons of the pound and its copies. From 1883 to 1933, the diminution amounted to 1 part in 5 million, which again was more than that suggested by the other source of comparison. Even disregarding the value obtained in 1846, which depends on the now discarded Kilogramme des Archives, the diminution of the pound between 1883 and 1933 is more than can be attributed to experimental error. For this period there is good evidence of the relative stability of the kilogram copies amongst themselves. Moreover, the quality of the material of which these were constructed is superior to that of the platinum pounds, which are of much earlier date. Furthermore, experiments made at the International Bureau, in which one of the kilogram standards of platinum-iridium was kept for a year in a stream of air at 100° C, demonstrated that the change in mass was almost negligible (Guillaume 1927). These considerations in combination give reasonable assurance that the kilogram can be regarded as a stable standard, and it must be concluded that the pound has changed. It is not proposed to speculate on possible causes of its changes.

---

* 1936, *Report by the Board of Trade on the comparisons of the Parliamentary Copies of the Imperial Standards.*

In connexion with the determinations of the relationship between the pound and the kilogram, it is of interest to remark that in the United States, where both the pound and the kilogram systems are in use, the pound is defined as bearing a fixed ratio to the international kilogram. The value accepted for this purpose was that determined in 1883.* The British pound, having diminished by 1 part in 5 million since 1883 in relation to the kilogram, is therefore smaller than the U.S. pound by that amount.

Secondary and other standards

The fundamental standards serve as the ultimate basis of weighings for purposes of trade, industry and scientific work, for which secondary and other standards are needed. Before considering how these are derived from the fundamental standards, a brief reference will be made to the materials available for their construction. For most purposes these should be non-magnetic, highly resistant to corrosion, and not too difficult to machine and polish.

On account of its high stability, platinum, preferably alloyed with a suitable proportion of iridium or similar constituent, is pre-eminently the most suitable material for the construction of a precision standard of mass. Owing to its high cost, however, it is not often used for this purpose, except for standards of specially high status, and also for small denominations for which the cost of the metal is not such a serious consideration. Rock crystal has sometimes been used for the construction of precision standards but it is readily electrified and is by no means easily discharged completely. A nickel-copper alloy containing these metals in approximately equal proportions has also been used, but not extensively. Stellite, an alloy containing cobalt, chromium, and a small proportion of tungsten or molybdenum, has also been used, but owing to its hardness the cost of making standards of this material is relatively high. Brass, without any protective coating, cannot, of course, be regarded as a suitable material for a precision standard; its use is still far more extensive than is justified by its properties, and should be restricted to weighings of an appropriately lower class. Plated brass is sometimes satisfactory, but the ordinary process of plating does not always ensure protection from chemical action. Of the metals used for plating, nickel and gold are not found very satisfactory, while platinum has been less extensively employed since the introduction of chromium and rhodium. During the last 25 years, some experience has been obtained in the use of nickel-chromium alloy and stainless steel. The commercial variety of the former, containing 80 % nickel and 20 % chromium, is normally satisfactorily non-magnetic and is being used with some success for secondary standards, and also for analytical weights. Its stability, though reasonably satisfactory, is not, however, comparable with that of platinum.

---

* The value actually accepted was that quoted in full in the original report, viz. 453·592 4277.

The stainless steel which has been in use in recent years is one which purports to be reasonably non-magnetic, being of the austenitic variety containing 18 % chromium and 8 % nickel. This type, however, has not been found sufficiently non-magnetic to give complete confidence in its use and should be superseded by a more reliable variety, e.g. that containing 25 % chromium and 20 % nickel, which is far superior to the 18/8 type in regard to non-magnetic properties. In spite of metallurgical progress there are not many suitable materials for precision work generally, and the choice for the time being is largely restricted to nickel-chromium, stainless steel and plated brass.

In order to determine and maintain a secondary standard pound or kilogram, occasional comparisons have to be made between it and a copy of the fundamental standard; these follow the usual procedure of double weighing, with strict allowance for air buoyancy. The density of the secondary standard, which is usually in the vicinity of 8 g./ml. differs widely from that of the platinum standard, which is about 21 g./ml. The differential buoyancy correction therefore amounts to 1 part in 10,000 of the mass being determined and accordingly requires to be assessed with great care. Assuming that both density values have been ascertained with sufficient precision, the accuracy of assessment of the buoyancy correction depends on the precision with which the density of the air is determined at the time of weighing. This is computed on the basis of the ascertained current values of the pressure, temperature and humidity of the air within the balance. Precautions are taken to introduce fresh air into the balance case but the chemical composition of the air is not actually determined. This may preclude the attainment of an accuracy finer than 1 part in 20 million in deriving a secondary standard of mass from the primary. Of course the buoyancy difficulty might be circumvented by carrying out the weighings *in vacuo*, but other sources of doubt would presumably be introduced depending for example on the variability of the adsorbed gases on the surfaces of the standards. Whether or not some suitable technique can eventually be evolved for overcoming such difficulties, it appears that the accuracy at present attained is ample for all ordinary purposes.

Further considerations of buoyancy and conventions for adjustment of weights

From the secondary standards are derived the more ordinary everyday standards of mass which for convenience will be referred to by the less explicit term 'weights'. In practice most weights, apart from the smallest denominations, are made of brass or of a material of a similar density. In general the range of densities of the materials in use rarely exceeds 7 % of their mean value. The differences in buoyancy between like denominations will therefore normally lie within 1 part in 100,000 of the mass. These differences, of course, vary somewhat with the air density but as the latter rarely varies by more than 10 % under most ordinary conditions of weighing, the buoyancy difference remains constant to within 1 part in a million of the mass. This enables a large amount of routine weighing to be made in practice without elaborate attention to buoyancy corrections. Thus it is the general practice to adjust weights on the basis of balancing in air a standard of mass of some conventional density representing brass, the material most commonly used for the construction of weights. This system of adjustment is used for purposes of trade, industry and science, where ordinary and not exceptional precision is required. For trade, the conventional density of brass associated with the standard in this country* is 8·143 g./ml.† Inspectors' standard weights and trade weights are adjusted to balance the standard in air of normal density; but as has been said before small changes in air density do not seriously affect the result.

Though buoyancy allowances are not normally made in trade transactions based on weight, a fine point arises in monetary transactions in converting from sterling (which is based on the pound) to the currency of a metric country (which is based on the kilogram). The conventional density 8·143 associated with the standard used for trade weighings in this country differs somewhat from that adopted abroad, and on this account the effective conversion from pounds to kilograms, or vice versa, will differ from the pure ratio of the respective units of mass. This difference, which is by no means negligible, appears to be a matter which ought to be referred to the International Committee of Weights and Measures, with a view to arriving at some agreed procedure.

Weights used in industry and science (as distinct from trade) are normally adjusted to balance in air a standard of mass of density 8·4 g./ml. and are calibrated on this basis in air whose density is sufficiently near the normal value not to affect the result appreciably. A typical example of this is a set of analytical weights, containing multiples and submultiples of a gram. These are usually so constituted that, once the value of the head weight is ascertained, the values of the remaining weights can readily be determined by intercomparisons among the weights of a set. Analytical weights are standardized on this basis at the N.P.L. to an accuracy of 1 part in a million for weights of 10 g. and larger, and to 0·01 mg. for smaller weights; this is quite adequate for the great majority of industrial and scientific purposes. As a result of this procedure the weights can, in effect, be treated as standards of mass of density 8·4 g./ml., even though their actual densities may differ slightly from this value. In any chemical weighings where allowances for buoyancy are of importance, these can be assessed so far as the weights are concerned by means of an inclusive correction for their buoyancy based on a density of 8·4 g./ml. This procedure is much simpler than that of calibrating weights according to their true masses, which would involve a knowledge of the actual density of each individual weight, and individual application of its buoyancy correction.

Receptacles for standards

In order to maintain the high stability of weights it is important that the receptacles in which they are stored should be appropriately designed and constructed, and should not contain any material which is likely to affect adversely the general surface condition of the standards. Manley (1933, 1935, 1945) has investigated the conditions of housing of weights in their cases and has studied the growth of films on the surfaces of weights of different materials under a variety of conditions of storage. His work, which is a substantial contribution to the knowledge of this subject, should be consulted in connexion with the designing of receptacles for weights.

---

* 1871, *Fifth Report of the Standards Commission*, Appendix XII, p. 170.

† There is no need here to discuss the reason for the adoption of this particular density value, which was introduced 100 years ago.

THE TEMPERATURE SCALE

BY J. A. HALL

The direct realization of the thermodynamic temperature scale involves the use of either a perfect heat engine or of a perfect gas, neither of which can be attained experimentally. The scale, therefore, must be reproduced by the use of one of the 'permanent' gases, the appropriate corrections to give the thermodynamic scale being deduced from the measured departures from the $PV = RT$ relationship under the particular experimental conditions to be used. The most generally employed gases are hydrogen, helium and nitrogen, and measurements are usually made in terms of pressure at constant volume. A gas thermometer, however, is an inconvenient instrument, and the experimental difficulties are considerable, increasing greatly at high temperatures.

The first step towards making the gas scale readily available was taken by Chappuis in 1888, when he compared four standard mercury-in-glass thermometers with his constant-volume hydrogen thermometer. The corrections needed to reduce the readings of these thermometers (which were used as absolute instruments to reproduce the 'mercury in verre dur' scale) to those of the constant-volume hydrogen scale were determined over the range 0–100° C to an accuracy of about $\pm 0{\cdot}005°$ C, and similar thermometers made from the same glass were distributed among various national standardizing laboratories and similar institutions. So was the 'échelle thermométrique normal' made generally available.

The next stage was the development of other methods of measuring temperature (e.g. the platinum resistance thermometer) which are capable of being read, in certain parts of the scale, to much higher accuracy than the gas thermometer itself. Such instruments, however, have not, in general, any theoretical connexion with the thermodynamic scale, and they can only be used to define it in so far as they have been accurately compared with a gas thermometer. They do not, therefore, improve the position as regards the accuracy with which we can define the thermodynamic scale. If, however, secondary instruments of this type are compared directly (or through some suitable intermediary) with a gas thermometer, they enable us to define the thermodynamic scale as accurately as the gas thermometer observations permit, while at the same time allowing us to reproduce a practical scale to a higher order of accuracy.

Such a practical scale is the International Temperature Scale, which was brought into being at the 7th Conference of Weights and Measures in 1927. As an example of the distinction between accuracy and reproducibility—a distinction which is always found to take a prominent place in temperature measurement—we may take the case of the boiling-point of sulphur. A study of all gas-thermometer data available at that time showed that the boiling-point lay between 444·50 and 444·60° C, and was almost certainly nearer the latter figure. By the use of a resistance thermometer, however, it is possible to reproduce the point to an accuracy of ± 0·005° C. On the International Temperature Scale, therefore, the sulphur point was fixed at 444·60° C by definition, though it was admitted that the second decimal figure was not accurately known. By this means we have a scale which can be reproduced to within ± 0·005° at different times and in different laboratories, and which is certainly in agreement with the thermodynamic scale to within ± 0·1°. Since that date, Blaisdell & Kaye (1941) have analysed all the published work on the subject, including recent experiments at the Massachusetts Institute of Technology, and recommend a revision of the value to 444·7° C.

Some idea of the way in which the accuracy of realization and the reproducibility of the thermodynamic centigrade scale vary over a wide range of temperatures is given in the following table.

Absolute accuracy of realization and reproducibility of the thermodynamic centigrade scale

Temperature °C	Accuracy of Realization	Reproducibility	Method
– 182·97	± 0·02	± 0·02	resistance thermometer
0	—	± 0·0003	"
100	—	± 0·003	"
444·6	± 0·1	± 0·005	"
1063	± 1	± 0·05	"
		± 0·1	thermocouple
		± 0·2	optical pyrometer
2000	± 6	± 2	"

The method by which the International Temperature Scale seeks to reproduce the thermodynamic centigrade scale is as follows. A range of fixed points which had been the subject of gas-thermometer measurement was selected. They were chosen so as to cover the temperature range from – 182·97 (boiling-point of liquid oxygen) to 1063° C (melting-point of gold), and thence up to the highest temperatures by the use of an agreed value of the radiation constant C_2. Between the fixed points the platinum-resistance thermometer and the platinum- 10 % rhodio-platinum thermocouple were chosen as interpolation instruments. The resistance thermometer was selected to cover the range – 182·97 to 660° C. It is calibrated by the usual quadratic law based on the ice, steam and sulphur points for temperatures above 0° C, while an extra calibration point at – 182·97° C provides the additional term required to define the gas scale accurately below zero.

The platinum thermocouple is used over the range 660 to 1063° C (melting-point of gold) and is calibrated by a quadratic equation through the gold, silver (960·5° C) and antimony (630·5° C) points. The antimony point is not a primary fixed point, and the value of the freezing-point of the particular specimen in use is determined by the resistance thermometer, thus ensuring continuity of the scale at this point. Unfortunately, however, the dividing point of the scale was fixed (for some reason which is not now clear) at 660° C. This has led to the curious result, pointed out by the Bureau of Standards at discussions on the revision of the scale which took place in 1939, that the freezing-point of pure aluminium cannot be measured on the International Temperature Scale. If a determination of the freezing-point is made with a resistance thermometer, the value found is just in excess of 660° C, and thus falls within the domain of the thermocouple. If, then, a thermocouple determination is made, the value is found to be just below 660° C and should therefore be measured by the resistance thermometer. The Bureau of Standards estimate that the discrepancy between the two parts of the scale is about 0·15° C, and in order to eliminate this difference, the 1939 meeting of the Comité Consultatif de Thermométrie* recommended that the scale should, in future, be divided at 630·5° C, the approximate value of the antimony point. The outbreak of war, however, made it impossible for the Conference on Weights and Measures to meet, and this recommendation has not been ratified.

Above the gold point extrapolation is by means of a radiation scale using monochromatic light, and taking 1063° C as the melting-point of gold and 1·432 as the value of C_2 in Wien's equation. These two constants are sufficient to define the scale.

There are one or two points which may be mentioned in amplification of the figures given in the table, which refer to the results obtained by using the apparatus and following the technique laid down in the specification of the International Temperature Scale.

It has been suggested (notably by the Physikalisch-Technische Reichsanstalt) that the triple point of water would provide a more reproducible fixed point than the normal freezing-point, but this proposal has not so far gained acceptance, as the view is held by some other workers that little is gained in accuracy to compensate for the loss in convenience. The value of the triple point is about 0·0098° C, of which about 0·007° is accounted for by the change in pressure and the remainder by the absence of dissolved air (the international specification calls for an air-saturated mixture of ice and water).

It will be seen from the table that it is the steam-point which, almost alone, limits the precision obtainable in the determination of the fundamental interval of a thermometer, though the extremely careful work carried out by Beattie and his collaborators (1935–6) at the Massachusetts Institute of Technology has somewhat altered this aspect of the problem. By using a closed system with an artificial atmosphere of helium in place of the usual open type of hypsometer they were able to obtain a fourfold improvement in reproducibility—somewhat better than

---

* 1939, *P.V. Com. int. Poids Mes.*

±0·001° C. By the use of the same technique they obtained a reproducibility of about ±0·002° C in the sulphur boiling-point.

Variations in the amount of deuterium oxide present in the water will have an effect on both the freezing- and boiling-points. As far as I have been able to ascertain, the maximum range of concentration which has been measured in water from natural sources is from 1 in 5600 to 1 in 8900. The melting- and boiling-points of D_2O are +3·8 and 101·42° C, so one arrives at an elevation of the freezing-point above that of pure H_2O varying from 0·00043 to 0·00068° C, with the corresponding figures of 0·00016 and 0·00025° C for the boiling-point. Evidently, then, heavy water need not worry us at the present stage of thermometric precision.

When the International Temperature Scale was first adopted, there was no very reliable evidence as to how closely the resistance thermometer scale agreed with the thermodynamic scale in the region between 0 and 100° C where extremely high precision (of the order of 0·001° C) is frequently required. A thorough intercomparison of resistance thermometers with a large number of mercury thermometers of the same construction as those used by Chappuis in his original intercomparison with the gas thermometer was therefore undertaken at the National Physical Laboratory (Hall 1930). This work established identity of the scales to within two or three thousandths of a degree over the range 0–50° C, while the discrepancies between 50 and 100° C certainly did not exceed 0·01° C and were probably a good deal less. It is not possible to be more precise, as Chappuis's observations above 50°C were confined to two temperatures only. Two pairs of observations were taken in chloroform vapour at 60° C and two pairs in alcohol vapour at 78° C. At 60° C the overall spread was 0·007° C, but there was a difference of 0·012° between the means of the two pairs of observations at 78° C. Giving weight to all these observations leads to a difference between the international scale and the thermodynamic scale of 0·007° at 80° C. One pair of the observations at 78° C is quite consistent with the resistance thermometer scale, and if Chappuis's equation is modified by ignoring the other pair of observations, then the difference is reduced to 0·003° C.

The question of the eventual elimination of the thermocouple as a standard interpolation instrument has been considered, and a certain amount of experimental work with the resistance thermometer has been done at the gold point. Moser (1930), at the Physikalisch-Technische Reichsanstalt, has obtained values ranging from 1061·92 to 1062·45° C, while some unpublished work carried out at the N.P.L. at about the same time led to provisional values in good agreement with the latter figure. The values found by Callendar (1899) and by Heycock & Neville (1895) when corrected to a sulphur boiling-point of 444·60° C (instead of 444·53° C), were 1061·2 and 1062·2° C. Bearing in mind the fact that the platinum available at that time was far from pure by modern standards, the agreement with the recent work is remarkable.

The internationally accepted value of 1063° C for the gold point is based on the gas-thermometer observations of Holborn & Day (1900) and of Day & Sosman (1911), which led to values of 1064 and 1062·4° C respectively. There is thus no justification for regarding the gold point as being known to a greater absolute accuracy than ± 1° C, and therefore no evidence that the quadratic extrapolation of the platinum-thermometer equation from the sulphur boiling-point does not hold up to the gold point. If 444·70° C is adopted as the value of the sulphur boiling-point, as suggested by Blaisdell & Kaye, the platinum-thermometer values for the gold point are increased by 0·7° C, and the agreement with the gas thermometer is even better. In view of the importance of the gold point as the start of the radiation scale, however, it is necessary to assign to it a definite numerical value, and a small correction term would be necessary to bring the readings of all platinum thermometers to the same value at the gold point, should the platinum thermometer become the recognized means of interpolation up to that temperature.

In using the radiation scale, one of the chief difficulties is to establish and verify the adequacy of suitable black-body conditions. Two methods are available for the optical pyrometer determination of the gold point (and also of the palladium point (1555° C)—a valuable secondary fixed point on the scale). One is the wire method, in which a short length of gold wire is used to complete the junction of a thermocouple inside the black-body enclosure. A second couple is used to control the rate of heating of the furnace and to maintain the furnace steady at the melting-point after the melting-point of the wire has been determined by observing the arrest in the thermo-e.m.f. and ultimate rupture of the circuit. This method, however, is not capable of as high precision as the use of a small ingot in which a black-body enclosure is immersed. The chief problem is to secure uniform temperature conditions in the furnace, and this has been solved in a variety of ways, such as by the use of multiple windings with or without the additional precaution of enclosing the crucible or black-body enclosure in a metal block. A reproducibility of ± 0·2° C can be obtained at this point. This is equivalent to ± 0·6 at 2000° C, but in practice the reproducibility at the higher temperature is only about ± 2° C. The additional uncertainty arises from possible errors in the value of the mean effective wavelength to be used in comparing the brightnesses of 2000 and 1063° C and in the angles of the two sectors which have to be used in turn in order to effect the big reduction in brightness (about 1200 to 1), which is necessary to compare the two temperatures.

A measurement of the absolute value of a temperature above 1063° C on the thermodynamic scale by optical means, however, depends on an absolute knowledge of the gold point and of the constant C_2 in Wien's (or Planck's) equation. An uncertainty of 1° at 1063° C is equivalent to 3° at 2000° C, and there is still the uncertainty in C_2 to be considered. In 1927 the value of 1·432 cm.deg. was internationally adopted. In 1939 the Comité Consultatif agreed that this figure was undoubtedly too low, but decided not to change to the value of 1·436, which was then considered the most probable, on the grounds that the change was hardly significant and that it was better to defer the matter until it seemed improbable that the value would soon need to be changed again. Since then Birge (1941) has derived from atomic constants a value of $1{\cdot}4384_8$, to which he assigns an accuracy of $\pm 0{\cdot}0003_4$. If, then, the uncertainty in C_2 be taken as ± 0·003, one arrives at an additional uncertainty of ± 3° C in temperature at 2000° C, making a total of

± 6° C. Here, again, the reproducibility attainable with the practical instrument is superior to the accuracy with which the scale is known.

The 1939 meeting of the Comité Consultatif de Thermométrie proposed that Planck's law should be substituted for Wien's law as the means of defining the radiation section of the International Temperature Scale, thus enabling it to reproduce the thermodynamic scale to higher temperatures. This proposal, which has not yet been ratified, would only affect a temperature of 2000° C by 0·3° or 3500° C by 0·7°. At 5000° C, however, the difference is 17° and at 6000° C, 50°.

Coming lastly to the most fundamental point, the absolute zero, the 1927 Conference did not attempt to relate the International Temperature Scale to the absolute scale, but only to the Centigrade thermodynamic scale. The position was reviewed by the Comité Consultatif in 1939 when the numerically lowest value considered was −273·144° C, obtained at Leiden, while the highest was a provisional value of −273·16 to −273·17° C relating to the work of Beattie, of which the final result had not then been published.

Considerable discussion took place as to the choice between −273·15 and −273·16° C, and finally the Committee passed a resolution that, in their opinion, 'according to the experiments completed at the present time', the most probable value was −273·15 ± 0·02° C. Since that date, Beattie (1941) has given a figure of −273·165 ± 0·015° C.

An interesting suggestion was placed before the 1939 meeting on behalf of the National Research Council of the United States, to the effect that the scale should be based on the absolute zero and the ice point instead of on the ice point and the steam point, the position of the absolute zero being fixed so as to give the steam point a value of 100·000° C at the present time. In this way a change in the value of the absolute zero would alter the numerical value of any point on the scale except the ice point: that is, the steam point would, in future, be just as liable to alteration in value as any other fixed point on the scale. While it was admitted that this plan would offer advantages to the worker at very low temperatures, it was agreed that, while the position of the absolute zero was still uncertain in the second decimal place, the inconvenience to the metrologist (who would be faced with the prospect of appreciable changes in the temperature known as 20° C) would of itself make it inadvisable to adopt the suggestion at the present time.

Finally, some mention may be made of the real need of industry for a temperature scale reproducible to the order of accuracy which has just been outlined. Mr Barrell has mentioned above the need for an accuracy of one or two thousandths of a degree at 20° C in order to secure adequate precision in the length of the yard and metre, and two other examples will show how the Laboratory is being pressed in other parts of the scale. At temperatures up to about 100° C, the oil technologists are regularly demanding an accuracy of 0·005° C for the determination of kinematic viscosity, while at high temperatures the steel industry is beginning to talk in terms of an accuracy of about 5° at 1600° C. The International Temperature Scale, therefore, has little enough in hand to satisfy modern industrial needs.

ELECTRICAL STANDARDS

BY L. HARTSHORN

Ideal units

To the theoretical physicist electrical standards are entirely unnecessary. A standard of some kind is of course essential if numerical values of the various concepts are to have any meaning in terms of practical experience, but if standards of length, mass and time have been established, the theoretical physicist can define the unit for any electrical property by simply writing down an equation. The theory of electromagnetism plays so important a part in electrical practice that the theoretical units derived in this way, by writing down the simplest possible sequence of equations that will cover the field under investigation, have long been recognized as an ideal system of units by both theoretical and practical men. Like many other ideals it was utterly impracticable when it was first conceived (by Weber in 1851 following Gauss in 1833) and for long afterwards, and it says much for the idealism of the electrical world that it followed the recommendations of the B.A. Committee of 1862, and adopted the system for all classes of work, theoretical and practical. It is curious when we look back to note that the unit first realized, the B.A. ohm of 1865, was in error by 1·3 %. The accuracy with which these theoretical units have been realized has, however, steadily improved down the years, and, although it has never kept pace with that of the most precise purely electrical work, the margin between the two is now so small that it is of no significance to anybody outside standardizing laboratories. An accuracy of 1 part per 100,000 has now been realized for all the more important electrical quantities. The present note outlines some of the more interesting metrological problems that have arisen in the course of this work.

Practical standards

The most precise purely electrical work mentioned above consists largely of measurements made by means of the Wheatstone bridge and its various modifications, and the potentiometer. These devices have now been brought to such a state of perfection that it is a fairly simple matter to compare resistances and standard cells with an accuracy of 1 part per million. It follows that with equipment of this kind, and various combinations of resistors and standard cells, measurements of resistance, voltage, current, power, quantity can all be made without much difficulty with an accuracy about ten times as great as that with which the theoretical units are known. Moreover, modern resistors and standard cells appear to be stable over a period of at least a few years with something like the same accuracy, although this is a point not easily substantiated. It follows that if the highest possible precision at any one time were the chief requirement, the best we could do would be to adopt a selected group of resistors and standard cells as primary standards, i.e. on the same footing as the metre kilogram, and secondly, to relate all other electrical measurements to them. This is, indeed, the course followed in the standardizing laboratories. Each national laboratory sends representative coils and cells, with their estimated values in theoretical units (or whatever other unit is called for in the current agreed international specification), to the International Bureau of Weights and Measures, where they are compared, and as a result of the comparison 'mean international units' are adopted with a precision of 1 part per million. In this way arbitrary units are accepted by the national laboratories for the purpose of securing the greatest possible degree of international uniformity, but the procedure ensures that the differences between these arbitrary units and the theoretical (or specified) unit are as small as it is possible to make them; for when assigning their values the national laboratories take into account all the available evidence concerning the theoretical or other specified units.

At the present time the specified units are the 'international units' defined by reference to a column of mercury and the silver voltameter, these units having been adopted by the International Conference held in London in 1908, contrary to the opinions of the President, Lord Rayleigh, because the Conference as a whole could not be persuaded that the theoretical units were really practicable.

Recent work has, however, shown that the theoretical units are now more practicable than the specified units, and there is little doubt that the international ohm, ampere, etc., will soon be dropped in favour of the theoretical ohm, ampere, etc. The object of the pioneers, viz. strict consistency between theory and practice, will then have been achieved as far as is possible, though the arbitrary element in the units employed at the standardizing laboratories is always likely to remain.

Relations between electrical and mechanical quantities

Returning now to the work that has proved beyond reasonable doubt that the theoretical units are practicable, note that the problem to be solved is to establish links between the mechanical quantities that are measurable by reference to the standards of mass, length and time, and the various electrical quantities. In theory, *the* link is energy which is common to the two systems, but unfortunately this quantity, in spite of its dominating importance in physical science, can only be measured by indirect and correspondingly inaccurate methods; and this link, introduced by Thomson (Kelvin) in 1862, is of no direct value in modern metrological practice.

The links that would be of most value in practice are those directly connecting the three fundamental quantities length, mass and time with some electrical quantity. The simplest of such relations is that between time and frequency, a property of alternating currents, but this barely takes us into the electrical world. Length is directly linked to capacitance and inductance. Of these only the inductance link has been established with the accuracy now required, and the unit of inductance is therefore the starting-point of the theoretical system. Several electrical quantities, including resistance, can be measured in terms of inductance and frequency by the technique of the a.c. bridge, but they are all what the electrical engineer calls 'dead' quantities, mere ratios of the 'live' quantities, and to complete the system we must have a link between a 'live' quantity and mass. There is no direct link with mass alone, but in the current balance there is a system in which current is related to force and the geometry of the circuits. It therefore gives a somewhat indirect link between current and mass, length and time, expressed by the equation

$$I^2 = mg/D \quad \text{or} \quad I^2\frac{dM}{dz} = mg.$$

where I denotes current, m mass, g gravitational acceleration, and D the dynamometer constant of the fixed and moving coils, which is a derivative of mutual inductance and directly linked with length.

These few links are sufficient to determine the whole system of electrical and magnetic quantities: standard resistors are calibrated in terms of mutual inductance and frequency, standard cells in terms of current and resistance, and everything else is measured by reference to those standards and the standard of inductance.

Mechanical to electrical transfer standards

The problem of establishing these links between the electrical and mechanical units consists essentially in constructing apparatus on which it is possible to make measurements of the highest precision of some electrical quantity and also of the mechanical quantities to which it is related by theory. It is necessary to satisfy at the same time the conditions for precise mechanical measurements, for precise electrical measurements, and those postulated in the theoretical equations. These are very severe conditions, and there is some excuse for those who hold that a system of units which was dependent on the realization of at least two such links was impracticable. For mechanical precision attention must be confined to materials that are sensibly rigid and of low temperature coefficient of expansion. For electrical precision only the best conductors and the best insulators must be used, and for theoretical precision there must be the closest possible approach to linear circuits of zero resistance in a medium that is perfectly non-conducting and non-magnetic.

As is well known three such links have been established at the N.P.L., the Campbell standard of mutual inductance, the Ayrton-Jones current balance, and the Lorenz apparatus. A detailed consideration of the way in which the difficulties have been met in each of these is beyond the scope of a short paper, but it may be noted that mutual inductance is a common factor in the three links, and it follows that the accuracy of the whole system of measurements turns very largely on the accuracy with which inductors of suitable size and form can be constructed and measured. The main features of this side of the work will therefore be noted by way of illustrating its general character.

The type of construction employed in all this work at the N.P.L. is that of the single-layer solenoid, wound with thin bare copper wire under tension on a marble cylinder. The tension in the wire ensures that the geometry of the coil is almost entirely determined by that of the marble, this material having been chosen for its rigidity, low temperature coefficient of expansion ($4\text{–}6 \times 10^{-6}$ per 1° C), and tolerably good insulating and machining properties. Lord Rayleigh recently showed that the dimensional stability of marble is far from perfect, and its insulating properties may become poor owing to its hygroscopic character, but nevertheless under suitable conditions it has given satisfactory results, surpassed only by the most recent American work with cylinders of pyrex glass, the working of which is, of course, far more laborious. Accurately defined radial dimensions are automatically obtained by winding on to a cylindrical surface, but the axial dimensions of the coils are equally important, and the precise axial location of the wire is far more difficult. It becomes essential to fix the position of every turn of the coil by winding the wire in helical grooves cut into the surface of the marble cylinder. The machining of these grooves with the necessary precision is one of the most important steps in realizing the units.

It is of the greatest importance that the finished coil shall have the highest possible degree of geometrical uniformity, not only because the theoretical formulae can only be applied to circuits of regular form, but also because only in such a case can the geometry of the circuit be completely specified by a finite number of measurements of length. In order to secure this uniformity, not only must the helical groove be uniform but the wire must be of uniform circular cross-section. One coil was found to show irregularities of diameter because the wire was of elliptical cross-section and the ellipse had rotated during the winding. In the most recent American work, the final drawing of the wire and winding of the coil were carried out in one operation, the wire passing through specially selected dies straight on to the cylinder.

Corrections for departure from theoretical conditions

Appreciable departures from uniformity are inevitable when dealing with a material like marble, and it is therefore necessary to make measurements of diameter and axial length for a very great number of pairs of points throughout the length of the helix, and to make allowance for the observed irregularities by making approximate calculations of the effect of each and applying them as 'corrections' to the theoretical formula for a uniform helix. The whole process of both measurement and calculation is very laborious, but quite straightforward: the various turns of the helix are regarded as equivalent to circles of a diameter and axial position equal to the mean measured value for the corresponding portion of the helix.

This non-uniformity of the helix is not the only feature in which some departure from the conditions postulated in the theory is inevitable, and in every instance the corresponding 'correction' must be estimated and applied. Thus the wire although thin (0·6 mm. diameter) is not a linear conductor. As a first approximation one naturally regards the current as concentrated in the axis of the wire, but in some cases this approximation is not close enough. In order to find a better it is necessary to know the distribution of current throughout the cross-section, and this is not known. If the wire is regarded as equivalent to a bundle of parallel filaments, allowance must be made for the fact that the outer ones will be longer than the inner ones, and therefore of higher resistance if the resistivity of the material is uniform over the cross-section of the wire. Thus the outer filaments must be credited with less current than the inner ones so that the equivalent single filament is not in the axis of the wire. But here again an unjustifiable assumption has been made: the wire is under tension and the inner filaments are therefore compressed by pressure on to the marble, while the outer ones are stretched; the assumption of uniform resistivity is not justified. The best one can do here is to calculate the correction on the most plausible assumptions. Fortunately, the uncertainty is not large enough to be of any real consequence, though this effect of variable resistivity is believed to affect the temperature coefficient of the best American inductors. Such corrections are less important in mutual inductors than in self-inductors, because in the mutual inductance the uncertainty is confined to large quantities, the distances between the primary and secondary, while in the self-inductor it appears also in much smaller quantities, e.g. distance between adjacent turns.

A correction must also be applied to allow for the fact that materials like marble and pyrex glass are not strictly non-magnetic. Marble is diamagnetic, its permeability being about 10 parts per million less than unity, but since the cylinder is always hollow the 'medium' is by no means all marble, and the 'correction' for permeability only amounts to less than -3 parts per million in practice.

The effect of the finite resistance of the coils must also be considered. A definite voltage is required to drive the current against this resistance, and the coil possesses an electric field, as well as the magnetic field, in terms of which the theoretical units are defined. This field may give rise to attractive forces in the current balance or to induced voltage in the Lorenz apparatus, or to 'capacitance effects' when an inductor is used with a.c. These effects can usually be made negligible by a suitable choice of the electrical conditions, e.g. the potential distribution with respect to earth, or the frequency of the a.c., which must be kept sufficiently low; 10 cyc./sec. is usual.

Conclusion

This very brief survey is sufficient to show the complexity of the problems that have arisen in attempts to approximate as closely as possible to an ideal system of electrical units. So many disturbing factors are discovered in any one piece of apparatus that one is always asking how many have been missed. It is a melancholy historical fact that the experimenter's estimate of his 'probable error' is almost invariably far less than his actual error proves to be. There is, however, one good criterion of accuracy. If observers in different countries get consistent results for the same quantities by different methods of measurement, then they must have realized the same system of units, and it is but reasonable to suppose it is the one they were both looking for.

All the determinations of the ohm and the ampere that have been published from 1870 to the present time are summarized in figures 3 and 4. They have, as far as possible, been reduced to common bases by reference to the surveys given by

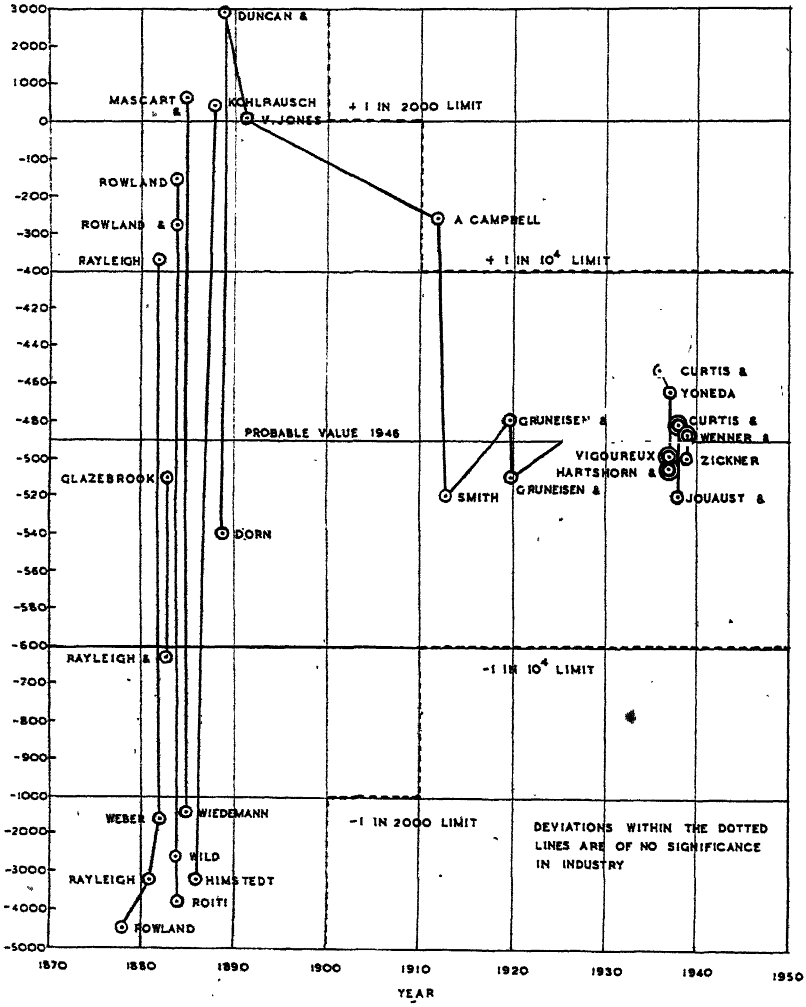


FIGURE 3. Determinations of the ohm. Only one author is named for each determination; the symbol '&' following a name means that there were collaborators. For further details see Smith (1922) and Curtis (1944). Two values are marked 'Gruneisen &'; they indicate an uncertainty in the value of the international ohm used by these observers.

Smith (1922) in Glazebrook's *Dictionary of Applied Physics*, vol. 2 and by Curtis (1944). These charts illustrate well for both the ohm and the ampere the steady approach to definite values that are consistent with all the observations made by all the methods within what may reasonably be regarded as their limits of error.

There is good reason to believe that the points in double rings obtained in the most recent measurements at the N.P.L. and N.B.S. represent the most accurate determinations, and it is seen that these deviate from the final value marked on the charts by little more than 1 part in 100,000. The maximum deviation of the three values for the ampere is only 4 parts per million, and of the four values for the ohm only 12 parts per million. It is exceedingly improbable that any of these results are significant to 1 part per million, and there is no reason to suppose that the ampere is known with greater certainty than the ohm. It may, however, be concluded that the uncertainty in both values is approximately 1 part per 100,000, or about one-tenth of the smallest amount that is significant in industrial measurements.

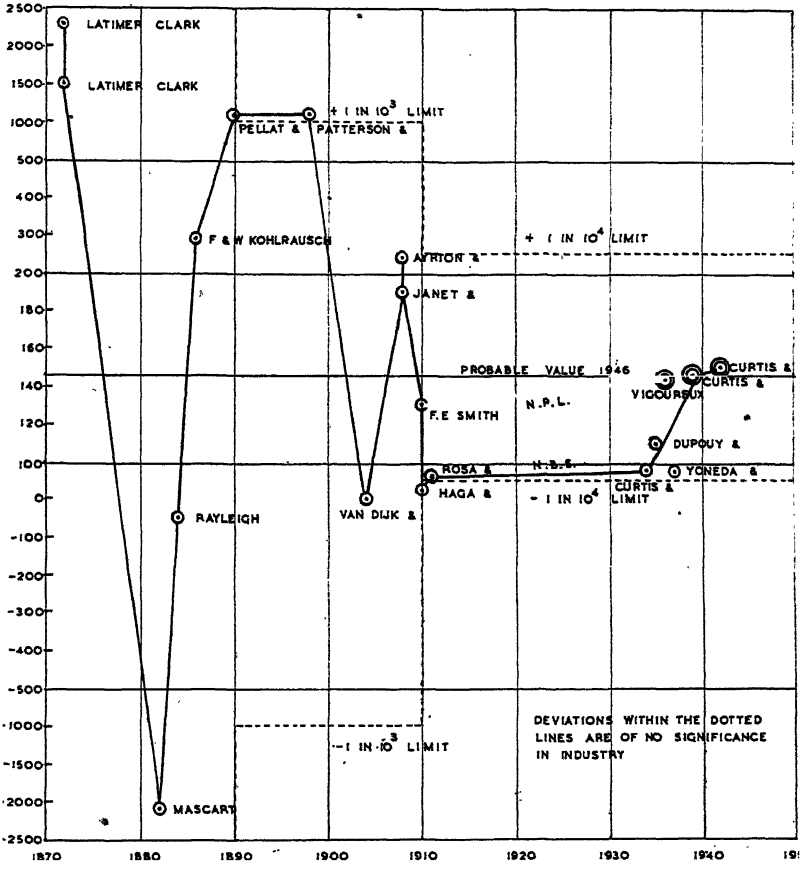


FIGURE 4. Determinations of the ampere. The symbol '&' following a name has the same meaning as in figure 3. For further details see Smith (1922) and Curtis (1944).

THE ACCELERATION DUE TO GRAVITY

BY J. S. CLARK

Most determinations of the acceleration due to gravity are relative measurements, the resulting values of g being made to depend on some one reliable determination of the absolute value. The relative values of g in such a system are in general more accurate than the computed absolute values.

The values of g at two stations are usually compared by swinging a pair of so-called 'invariable' pendulums in each of the stations in turn. At each station, the two pendulums, which usually have semi-periods of about half a second, are set swinging in opposite phases so as to eliminate the effect of the motion of the support on which the pendulums are swung. Other methods of comparing the values of g at two stations have been used, but the above method is the one most generally employed, and is capable of giving an accuracy of about one part in a million.

Only a very few absolute determinations of g have been made in recent times, and the method used in each case was essentially that of the Kater 'reversible' pendulum, wherein the pendulum was adjusted until its period of oscillation about one of two alternative points of support was made very closely equal to its period about the other point of support when the pendulum was reversed end for end. The points of support may be knife-edges attached to the pendulum (as in the original Kater pendulum) or planes attached to the pendulum, the knife-edge being on the support. In either case, the length of the equivalent simple pendulum depends on the distance between the two points of support.

The accuracy of determination of the absolute value of g by the reversible pendulum method is inferior to that of relative determinations because, in addition to the corrections which also have to be applied to the observed periods of the invariable pendulums, there are a number of corrections peculiar to the reversible pendulum which are difficult to estimate with precision. These arise from the flexure of the pendulum itself, and from the fact that the pendulum does not oscillate about a geometrical point, but swings about a material knife-edge in contact with a plane surface. The knife-edge usually has a small, but finite, radius of curvature, and it undoubtedly suffers deformation under load. It is likely that the principal cause of uncertainty in the results obtained by means of reversible pendulums is due to the use of knife-edges.

Since 1909, the accepted international base for gravity measurements has been Potsdam. The accepted value, 981·274 cm./sec./sec., is the result of an absolute determination made by Kühnen & Furtwängler (1906) in the 'Pendelsaal' at the Königliche Preussische Geodätische Institut in Potsdam, and described by them in 1906. It corresponds with the value 981·275 cm./sec./sec. for the 'Ostkeller' of the Geodetic Institute.

Recent absolute determinations of g indicate that the Potsdam value is too large.

In 1936, Heyl & Cook published the result of an absolute determination made at the National Bureau of Standards, Washington, viz. 980·080 ± 0·003 cm./sec./sec. This result is 0·020 cm./sec./sec. lower than the value derived from the Potsdam value by means of a direct connexion between the two stations made in 1933 by Brown (1936) of the United States Coast and Geodetic Survey.

In 1939, the author published the result of an absolute determination made at the N.P.L., viz. $981{\cdot}181_5 \pm 0{\cdot}001_5$ cm./sec./sec. This result is 0·0138 cm./sec./sec. lower than the value derived from the Potsdam value by means of number of transfer measurements adjusted by Bullard & Jolly in 1937.

In 1940, Browne & Bullard published the results of direct intercomparisons made between the N.P.L., Teddington, and the N.B.S., Washington. These relative measurements gave the difference

$$g(\text{Teddington}) - g(\text{Washington}) = 1{\cdot}0969 \text{ cm./sec./sec.},$$

the difference between the two absolute determinations being 1·1015 cm./sec./sec.

In 1942, Dryden published the results of his analysis of the corrections which Kühnen & Furtwängler applied in the course of their original Potsdam determination. He suggested that the systematic errors present in the individual Potsdam determinations arose from other causes than those discussed by them, and that in the absence of further knowledge these errors should be considered as of a random nature, in which case a straight mean value, about 12 parts in a million lower than the value adopted by Kühnen & Furtwängler, would better represent the true value of g at Potsdam.

The following figures summarize the above absolute and relative determinations:

		Residuals
(A) Potsdam (amended by Dryden)	= 981·262	$0{\cdot}002_5$
(B) N.B.S. (Heyl & Cook)	= 980·080	0·003
(C) N.P.L. (Clark)	= $981{\cdot}181_5$	$0{\cdot}000_5$
A–B (E. J. Brown)	= +1·174	$0{\cdot}002_5$
B–C (Browne & Bullard)	= $-1{\cdot}096_9$	$0{\cdot}001_1$
C–A (Bullard & Jolly 1936)	= $-0{\cdot}078_7$	$0{\cdot}000_2$

The least-square solution of these six results gives

A (Potsdam) = $981{\cdot}259_5$ cm./sec./sec.
B (N.B.S.) = 980·083
C (N.P.L.) = 981·181

from which the residuals quoted above have been calculated. Using the original (Kühnen & Furtwängler) Potsdam value 981·274, the least-square solution gives

		Residuals
Potsdam	$981{\cdot}265_5$	$0{\cdot}008_5$
N.B.S.	980·086	0·006
N.P.L.	981·184	$0{\cdot}002_5$

No attempt has been made to weight the individual results contributing to these calculations. There would be obvious difficulty in doing so, and the smallness of the residuals indicates that no important change in the calculated values would be likely to result from any reasonable degree of weighting that might be suggested.

According to Dryden (1942), the Sub-Committee on Gravity of the National Research Council Committee on Fundamental Physical Constants has recommended a reduction of values in the Potsdam system by 17 parts in a million in those cases where the most precise absolute values are desired. The above analysis, however, supports the view that values in the Potsdam system should all be reduced by $0{\cdot}014_5$ cm./sec./sec. to give the closest approximation to the absolute values.

It seems unlikely that any future absolute measurements made with reversible pendulums will lead to appreciable improvement in accuracy, despite the very high precision with which measurements of length and time can now be made. There is, however, at least one other possible method of determining g absolutely, which does not involve the corrections associated with the use of a pendulum and knife-edges, and that is by observing the time of free fall of a body in vacuo.

Standard gravity

For certain purposes, e.g. the measurement of barometric height, which varies with the acceleration due to gravity, the need for relating the results to some gravity standard has long been recognized. Originally this standard was taken to be gravity in latitude 45°, at mean sea-level. No particular value of gravity was at first associated with this standard, and the reduction of results on this basis was made by means of simple theoretical formulae providing for changes of gravity with latitude and altitude. Later, when the need arose for associating a particular value with the gravity standard, the value 980·665 cm./sec./sec. was accepted by the Third International Conference for Weights and Measures in 1901.* This value was subsequently found to be somewhat too high, but it was deliberately retained by the Fifth International Conference in 1913 as the value of standard gravity for reference purposes, and was thereby given a conventional status distinct from the original conception of standard gravity. The same value was used in the definition of the standard atmospheric pressure to which boiling-points on the International Temperature Scale are referred.†

A different view of standard gravity, however, was taken in regard to meteorological work under the auspices of the International Meteorological Committee, and the geophysical basis, viz. gravity in latitude 45° at mean sea-level, was retained. Associated with this is the rounded value 980·62 cm./sec./sec.‡

---

* 1901, *C.R. 3rd Conf. Gén. Poids Mes.*

† 1927, *C.R. 7th Conf. Gén. Poids Mes.*; 1933, *C.R. 8th Conf. Gén. Poids Mes.*

‡ *Organisation Météorologique Internationale*, no. 45, Comm. Mét. int. Proc.-verb. Berlin, 1939. Lausanne, 1941, p. 75.

The concurrent use of two different conceptions of standard gravity, each supported by the authority of an international organization, is unsatisfactory and is a cause of confusion which should be eliminated. The whole subject of gravity in its bearing on geodesy, physics, meteorology, etc., needs to be reviewed with due regard to changes which now appear necessary in the absolute basis of gravity.

THE BAROMETRIC STANDARD

By F. A. Gould

Fundamental considerations

In connexion with the establishment of some of the standards discussed in other papers of this series, an accurate knowledge of atmospheric pressure is necessary. Examples of this may be found in the determination of the steam point and other basic points on the international temperature scale; also in the determination of air density in connexion with certain precision weighings. The mercury barometer provides the most accurate means of determining atmospheric pressure. Barometer scales were originally graduated in units of length, and the process of determining pressure consisted, in principle, of two stages, viz.

(i) the determination of the true height of the mercury column from the scale reading, and

(ii) the translation of that height into a conventional pressure unit.

The first stage is largely a matter of instrumental convention and technique, which will be considered in a later section of this paper. In the second stage it is evident that if the density of mercury and the value of gravity in the vicinity of the barometer are known, the pressure can be expressed in absolute units, e.g. dynes/sq.cm., in accordance with the fundamental hydrostatic equation: $p = h\rho g$. Further, the scale can be graduated in units of length so chosen that the reading taken under specified conditions of temperature and gravity is identical with the pressure expressed in absolute units. In effect, such a scale would represent the result of multiplication of the three quantities, h, ρ and g, in the hydrostatic equation. This principle is adopted in the graduation of the 'millibar' scale (1 millibar = 1000 dynes/sq.cm.).

Apart from the millibar, which is of comparatively recent introduction, the conventional pressure unit adopted is a column of mercury of unit length (inch or millimetre) measured at 32° F (0° C) and subject to standard gravity.

The use of this reference temperature follows a long-established practice, but there are, unfortunately, two conceptions of standard gravity, each supported by the authority of an international organization.* The International Meteorological Committee retain the original conception of standard gravity, i.e. gravity at mean sea-level in latitude 45°, and now associate with it the value 980·62 cm./sec./sec.

---

* See 'The acceleration due to gravity' in this series of papers.

On the other hand, the International Conference for Weights and Measures has defined standard gravity as being equal to 980·665 cm./sec./sec., regarding it as a conventional reference standard distinct from the original geophysical conception. This value is used by physicists, chemists and engineers generally in work of high precision carried out in this country. In addition, the International Conference for Weights and Measures* has defined standard atmospheric pressure as the pressure exerted by a column of mercury, 760 mm. high, and of density 13·5951 g./c.c., subject to an acceleration of 980·665 cm./sec./sec. due to gravity. This definition is followed in precision pressure measurements made in the determination of physical standards and constants.

Fundamental standard barometers

(*a*) *Early history.* For many years the standardization of barometers was carried out at Kew Observatory where there were two mercury barometers, both having tubes of large diameter and designed to be read by cathetometer. These were installed between 1855 and 1860, and served as the fundamental standards of the Observatory. In 1912, when barometer testing was transferred to the N.P.L., the two standards remained at Kew, and a Fortin type barometer, having a tube of internal diameter 19 mm., was used as the principal reference standard at the Laboratory pending the design and construction of a new primary standard barometer, to be installed at Teddington. The Fortin barometer was compared with the Kew standards before being put into service and was subsequently standardized occasionally from first principles at Teddington. It was later compared with the new primary standard barometer, which was installed at the N.P.L. in 1931.

During the interval, 1912–31, the Fortin barometer remained in exceptionally good condition without need of overhaul, and no appreciable change in its error was found. A high degree of accuracy was attained in the maintenance of a barometric standard in rather difficult circumstances, and no appreciable discontinuity was introduced by the use of different standard barometers in 1912 and 1931.

(*b*) *Primary Standard Barometer at the N.P.L.* In this barometer a departure has been made from the usual conception of an instrument based on the use of a glass barometer tube containing the mercury column (see figure 5, plate 6). Instead of glass, a block of stainless steel is used, suitably bored to take the barometric column of mercury, which is of large diameter at the top and bottom, where the mercury surfaces are viewed through optically flat and parallel glass windows sealed to the body of the barometer. Provision is made for the vacuum space above the mercury column to be renewed, as required, by a suitable pumping system. The measurement of the length of the column is referred to a line-standard of length mounted vertically near the column, and is made by means of a vertical comparator which carries a pair of micrometer microscopes disposed one above the other. The comparator can be traversed so that the microscopes are brought opposite either to the mercury surfaces or to the line-standard as required.

---

* 1927 and 1933, *C.R. Conf. Gén. Poids Mes.*

The steel body of the barometer is also bored vertically to take a mercury thermometer having a bulb which is adjacent to and extends over the whole length of the barometric column. This thermometer, being immersed in mercury and surrounded by the same mass of steel as the barometric column, records very closely the mean temperature of the latter. The reader is referred to the original account (Sears & Clark 1933) of this instrument for further details, including the optical system for reading the position of each mercury surface.

(*c*) *Reduction of measurements.* The first stage in the determination of pressure with this barometer is the measurement of the true height of the mercury column at the prevailing temperature, which can be controlled thermostatically. The height is then corrected to give the equivalent height of the column at 0° C. The coefficient of thermal expansion of mercury which is used for this purpose is sufficiently well established to provide for this correction without any appreciable loss in accuracy.* Further, to express the pressure in terms of a column of mercury at 0° C and under standard gravity 980·665 cm./sec./sec., the height of the column at 0° C is multiplied by the ratio of gravity at the N.P.L. to standard gravity, i.e. by $\frac{981\cdot181_5}{980\cdot665}$.

Alternatively, if the pressure is desired in dynes/sq.cm., the height of the column, reduced to 0° C but subject to gravity at the N.P.L., is multiplied by the density of mercury at 0° C, which is taken to be 13·5951 g./c.c., and also by the value of gravity at the N.P.L., viz. $981\cdot181_5$ cm./sec./sec.

(*d*) *Accuracy.* It is estimated that under reasonably steady conditions the barometric height can be determined to an order of accuracy of ± 0·002 mm., viz. to 1 part in 400,000 approximately. The accuracy associated with the absolute determination of gravity made at the N.P.L. in 1937 is ± 0·0015 cm./sec./sec., (Clark 1939), corresponding to 1 part in 700,000. So far, however, the density of the mercury used in this barometer has not been ascertained, and this is the chief cause of uncertainty in the determination of pressure. The ordinary method of purifying mercury by distillation under reduced pressure results in a partial separation of its isotopes. For this reason the densities of different specimens of mercury vary slightly according to their history. The evidence available from relative and absolute density determinations suggests that the density of mercury should not be regarded as constant to closer than ± 1 part in about 150,000 of the mean value 13·5951 g./c.c. assumed. When due allowance for all sources of error has been made, it is considered that an accuracy of ± 1 part in 100,000 is attained in the determination of the atmospheric pressure with the Primary Standard Barometer.

Secondary standard and other mercury barometers

(*a*) *Usual types.* There are numerous varieties of design of mercury barometers, and these vary considerably in different countries. It is proposed here to refer briefly to two general types of barometers which have long been in use in this country, viz. the Fortin and Kew types. In the Fortin barometer, settings are made at the lower and upper mercury surfaces. For the former, a fiducial point fixed in the ceiling of the cistern serves as the zero to which the mercury surface is brought through the agency of Fortin's leather bag and an adjusting screw. The level of the top of the barometric column is determined by a setting with a cursor carrying a vernier which is read against the scale of the instrument. In the Kew barometer only a single setting is required, viz. at the top of the column, this being effected as in the Fortin barometer. To compensate for changes in level of the mercury in the cistern, the scale is contracted uniformly. The amount of contraction depends on the diameters of the tube and cistern.

---

* See, for example, *Proc. Amer. Acad. Arts Sci.* **74**, 371, 1941.

Of the two types, the Fortin barometer is often preferred as an observatory or laboratory standard, but the Kew barometer is much more widely used for general purposes owing to its greater simplicity of design. For high accuracy, the tubes of barometers should be of large diameter. Small tubes are used where portability is of greater importance.

(*b*) *Scale units and conventions employed.* The scales of barometers of the Fortin and Kew types are normally graduated on brass. The inch scale of a mercury barometer is nominally correct at the temperature 62° F associated with the definition of the Imperial Standard Yard. Under this convention the scale measures correctly in inches the current length of the mercury column when the temperature of the barometer is 62° F.* The convention is applied to all inch barometers, irrespective of their type. Thus the readings of all inch barometers should agree at 62° F.

In conformity with the temperature 0° C at which the international metre is defined, the millimetre scale of a mercury barometer measures correctly in millimetres the length of the mercury column when the temperature of the barometer is 0° C. Accordingly, the readings of all millimetre barometers should agree at 0° C.

The millibar scale is employed almost exclusively for meteorological purposes, and the current conditions underlying its use were formulated by the Meteorological Office in 1914. The scale is so graduated that its reading gives the pressure correctly in millibars when the barometer is at a temperature of 285° A (12° C),† and is at mean sea-level in latitude 45°.

It is unfortunate that the conventions governing the graduation of the inch, metric and millibar scales lack uniformity. This causes some confusion in the correlation of corresponding readings on the scales and also in their reduction; and renders it difficult to prescribe for a fourth scale, viz. a scale of altitudes, which is used on one type of long-range barometer. Consideration should therefore be given to the possibility of simplifying the current system of scale conventions.

---

* Effect of capillary action. In practice, barometers are adjusted to agree ultimately with a fundamental standard instrument in which the effect of capillary action is negligible. This adjustment involves a slight displacement of the 'zero' of the scale, so that the reading corresponds, not to the length of the actual mercury column, but to the length which the column would have in the absence of capillarity.

† For meteorological purposes, temperatures expressed in ° A are derived from temperatures measured in ° C by adding 273°.

(*c*) *The testing of mercury barometers.* The testing of mercury barometers is undertaken at the N.P.L., where they are compared directly with a secondary standard of the Laboratory. The standard used for this purpose, a long-range instrument of appropriate design, is compared as required with the Primary Standard Barometer at the prevailing atmospheric pressure. This comparison is supplemented by others made from first principles, using as basis a barometric column in a large-bore tube in which the mercury levels are read with a cathetometer. It is estimated that an accuracy of ± 0·03 mm. is attained in the determination of the errors of the secondary standard.

In the routine tests, mercury barometers are set up in a large airtight chamber in which five instruments of ordinary design are compared with the secondary standard at the same temperature and pressure. The pressure in the chamber is set and controlled barostatically at any value in the range required, and the settings on the barometers are effected by suitable controls from outside the chamber, enabling the barometers to be read in the usual manner by means of their cursors and verniers. An accuracy of the order ± 0·1 mm. is attained in the standardization of barometers whose design admits of this precision. In many instruments, particularly those having tubes of small diameter in which the amount of capillary action is both considerable and variable, the accuracy attainable cannot be guaranteed to closer than ± 0·2 mm., or even ± 0·3 mm.

(*d*) *Reduction of barometric readings.* The term 'reduction' relating to barometric readings is normally understood to cover the corrections which have to be applied to the readings in order to express the pressure in some conventional form, either in absolute pressure units or in terms of a column of mercury measured under specified conditions of temperature and gravity. If the latter form of expression is chosen, it is important that these conditions should be clearly stated. For example, the definition of the gallon lacks precision on account of the omission to state the conditions.*

In the more precise determinations the reduction is made from first principles with due allowance for errors in measuring the height of the column, for its temperature, and for the value of gravity at the station where the barometer is read. So far as the temperature reduction is concerned, the Fortin and Kew types of barometer have a somewhat different temperature coefficient. To avoid misunderstanding, the appropriate reduction formulae for the respective types are given in the certificates issued for barometers tested at the N.P.L. unless provision for effecting the reduction is otherwise made. If the value of gravity at the station is not known it can be assessed to a reasonably good approximation, in general, by reference to *International Critical Tables*.† With this information the equivalent height of the column under 'standard' gravity can readily be computed.

---

* See 'The measurement of volume' in this series of papers.

† *International Critical Tables*, 1, 401. New York and London: McGraw-Hill.

For ordinary purposes published tables of temperature and gravity reductions are available.* These lack satisfactory provision for the temperature reduction of Kew barometer readings in inches and millimetres, and are confusing in regard to the two different conceptions of standard gravity employed. More comprehensive reduction tables for inch and millimetre barometers are needed, but before these can be prepared the whole subject of gravity values needs reconsideration with due regard to changes which now appear necessary in the absolute basis of measurement of gravity.†

Note. The scope of this paper has been restricted to determinations of pressure covering broadly the whole natural range of atmospheric pressures and based on the mercury barometer as standard. For the measurement of pressures of a higher order of magnitude, pressure gauges are ordinarily employed. These are standardized by determining the pressure applied by a piston loaded with known weights. Pressure gauges of all ranges up to 12 tons/sq.in. are tested on this basis at the N.P.L. to an accuracy lying within ± 1 part in 2000.

THE MEASUREMENT OF VOLUME

BY VERNEY STOTT

Units of volume

There are two kinds of units of volume, those directly derived from the corresponding unit of length, and those defined in terms of the volume occupied, under specified conditions, by a definite mass of water. In any system of units, once the fundamental unit of length has been defined, a hypothetical cube with edges of unit length clearly becomes a unit of volume. For practical purposes, however, hollow vessels are very convenient, and have long been used for measuring volumes of liquid, but to determine their capacity from linear measurements is not nearly so simple as to determine the weight of water required to fill them. Consequently units of volume came into existence defined in terms of the space occupied by a definite mass of water under specified conditions. The two types of unit are defined independently and the relation between them is a matter of experimental determination.

In the older systems of units there is no simple numerical relation between the two types. For example, the number of gallons in a cubic foot is not a whole number. In the metric system an attempt was made to link the two types of unit in a simple manner through the definition of the kilogram, and was achieved to a high degree of accuracy, but, as will be seen in the next section, even in the metric system the two types of unit are now defined separately and the relation between them is one based on experiment.

---

* *The Meteorological Observer's Handbook*, London: Meteorological Office; *Smithsonian Meteorological Tables*, Washington: Smithsonian Institute.

† See 'The acceleration due to gravity' in this series of papers.

Metric units

The founders of the metric system sought to ensure a simple numerical relationship between the two types of unit by defining the kilogram as the mass of a quantity of water occupying 1 dm.^{3} at its temperature of maximum density. Since, however, a cubic decimetre of water is manifestly unsuitable as a standard of mass, practical realization of the object in view necessitated making a standard of mass, of suitable material, strictly conforming to the above definition of the kilogram. The original Kilogramme des Archives was accepted as fulfilling this condition, and the litre, the volume of a kilogram of water at its temperature of maximum density, was regarded as equal to the cubic decimetre. Subsequently, however, the original definition of the kilogram was abandoned because doubt had arisen as to whether the Kilogramme des Archives did conform precisely to that definition, and the kilogram is now simply the mass of the International Prototype Kilogram.

The litre was separately defined* as being the volume occupied by a mass of 1 kg. of pure water at its temperature of maximum density and under a pressure of 1 atm. The reference to atmospheric pressure was introduced because water is slightly compressible. For pressures in the neighbourhood of 1 atm. and at ordinary room temperature a change in pressure of 1 mm. of mercury causes a change in volume of 0·000066 ml./l. Since the litre is defined on a mass basis, the buoyancy effect of the air will always result in the apparent weight in air of 1 l. of water being less than 1 kg. For example, against weights of density 8·4 g./ml. and in air of density 0·0012 g./ml., the weight of the quantity of water of temperature 4° C which occupies 1 l. at that temperature (and hence has a mass of 1 kg.) is 998·943 g. At higher temperatures the weight will be still less owing to the decrease in density of the water. For example, under the same conditions of weighing as above, the weight of a quantity of water of temperature 20° C and occupying 1 l. at that temperature is 997·176 g.

The present definitions of the kilogram and the litre contain no reference to the cubic decimetre, and when the litre was redefined arrangements were made for the accurate determination of the relation between the litre and the cubic decimetre. Three separate determinations were made.† The principle of each method was to determine by linear measurements the volumes of bodies of simple geometrical form and then to determine the mass of water displaced by the bodies by weighing them first in air and then in water. The methods differed only in the means employed to determine the linear dimensions. Guillaume used three hollow bronze cylinders, closed at the ends, of volumes approximately 780, 1300 and 2000 cm.^{3} and having diameters approximately 10, 12 and 14 cm. respectively and determined their dimensions by mechanical means. Chappuis used three glass cubes of 4, 5 and 6 cm. edge and employed a Michelson interferometer to measure the distances between the faces. The third determination was made by De Lepinay, Buisson and

---

* See *Trav. Bur. Int. Poids Mes.* **12** (1902).

† See *Trav. Bur. Int. Poids Mes.* **14** (1910).

Benoit using two quartz cubes of 4 and ·5 cm. edge. They also utilized an interference method to determine the separation of the faces of the cubes but one differing from that employed by Chappuis. The final means of the individual determinations were:

using bronze cylinders 1 l. = 1000·029 cm.^{3},
using glass cubes 1 l. = 1000·026 cm.^{3},
using quartz cubes 1 l. = 1000·027 cm.^{3},

and the value finally accepted on the basis of all the results was

1 l. = 1000·027 cm.^{3}

It is a tribute to the skill with which the Kilogramme des Archives was constructed in the eighteenth century that the litre approximates so closely to the cubic decimetre.

The difference in magnitude between the litre and the cubic decimetre and consequently between the millilitre and the cubic centimetre may often be ignored in comparison with the accuracy desired in a determination, but the distinction between the units should be clearly borne in mind and the difference taken into account when necessary. For example, densities expressed in g./ml. will differ numerically by approximately 3 parts in 100,000 from the same densities expressed in g./cm.^{3}, and so when densities of liquids are given to five decimal places the results expressed in g./ml. and those expressed in g./cm.^{3} will differ significantly in the fifth place.

To avoid much of the confusion which exists at present, it would be an advantage to express density—mass per unit volume—in g./ml. on all occasions, not only because of the difference just noted but also as affording a much better basis than specific gravity. This view is gaining acceptance; for example, density expressed in g./ml. is used throughout the *International Critical Tables*, and British Standard hydrometers are adjusted to indicate density in g./ml. at 20° C.

At one time confusion existed with volumetric glassware because some vessels were marked 'c.c.' although the unit used was the volume of a quantity of water having an apparent weight in air of 1 g. It has been stated previously that 1 l. of water of temperature 20° C has a weight in air of 997·176 g., so that a '1000 c.c.' vessel containing water of temperature 20° C having a weight in air of 1 kg. would be nearly 3 ml. greater in capacity than one containing either 1000 cm.^{3} or 1000 ml. at 20° C. Volumetric glassware, however, is now calibrated in terms of the millilitre, marked 'ml.' and also with the standard temperature 20° C; with such glassware no confusion can arise.

Imperial units

The gallon is the same type of unit as the litre. It is the volume of ten Imperial standard pounds of distilled water weighed in air 'against brass weights with the water and the air at the temperature of sixty-two degrees on Fahrenheit's thermometer and with the barometer at 30 in.' (§ 15, Weights and Measures Act, 1878).

The definition does not give precisely the conditions of weighing, and the Board of Trade assumes values of 0·00121698 g./ml. for the density of air and 8·143 g./ml. for the density of the brass weights to define the conditions under which 1 gal. of water of temperature 62° F has a weight of 10 lb. On this basis the gallon is equivalent to 4·54596 l. For conversion to units based on the Imperial units of length the relation 1 gal. = 277·42 cu.in. may be used.*

Heavy water

Prior to the discovery of isotopes, distilled water was regarded as a perfectly definite and homogeneous substance, and this assumption is implicit in the definitions of the litre and of the gallon. The average ratio of D_2 to H_2 in water is 1 to 6500, and if a sample of water contained D_2 in this ratio the complete removal of the D_2 would decrease the density by 17 parts in a million (Dorsey 1940). The variation in the D_2 content of ordinary distilled water is considerably less than that corresponding to the complete removal of D_2, and the lack of precision in the definition of the litre, in the light of the discovery of isotopes, is not likely to cause any practical difficulty in the measurement of volume.

Values of the density of water have, however, been published giving seven significant figures, e.g. to take only the values at one particular temperature from the tables, 0·9982336 g./ml. at 20° C (Chappuis as recalculated by the National Bureau of Standards (Tilton & Taylor 1937)) and 0·9982303 g./ml. at 20° C (Thiesen, Scheel & Dieselhorst 1900). The difference between these two values exceeds 3 parts in a million. The density of 'pure' water is therefore not known accurately to seven places of decimals, and the isotopic constitution would have to be taken into account in any attempt to establish the density to this degree of accuracy.

Volumetric measurements

Measurements of volume at the National Physical Laboratory have been mainly concerned with the verification of volumetric glassware. There is a great variety of such apparatus, some of which is limited in use to specific purposes, e.g. butyrometers, haemacytometer dilution pipettes, etc. The following details of the accuracy of certification offered for certain simple types of apparatus in general use serve to indicate the precision of measurement attainable in practice.

Type of Apparatus	Total Capacity (ml.)	Accuracy of Certification (v) (ml.)	v as Fraction of Total Capacity
measuring flask	1000	0·1	1 in 10,000
measuring flask	100	0·02	1 in 5000
pipette	100	0·02	1 in 5000
pipette	10	0·01	1 in 1000
pipette	1	0·005	1 in 200

---

* The present U.S.A. gallon derives from an old 'wine gallon' and is defined as being 231 cu.in.—inches being defined in the U.S.A. by the relation 1 m. = 39·370000 in.

The accuracy is lower for the smaller capacities than for the larger ones. The development of micro-analysis has created a demand for more accurate measurement of small volumes of liquid. A syringe-pipette recently designed at the Laboratory, in which liquid is expelled from a glass barrel by means of a well-fitting piston working between stops, enables an accuracy of 1 part in 1000 to be obtained on volumes of 1 ml. An accuracy approaching 1 part in 5000 is obtainable with a micrometer operated burette of similar capacity recently made at the Laboratory.

THE PHOTOMETRIC UNITS AND STANDARDS

From the Light Division, National Physical Laboratory

Historical

The candle was legalized by the Metropolitan Gas Act of 1860 which referred to the testing of gas in the metropolitan area. Apparently no unit was ever legalized for any other part of the country or for any other purpose. It had many disadvantages, and the Gas Referees who, under the Board of Trade, were responsible for gas testing under the Act, in their published 'Notification' for 1898 discarded the sperm candle in favour of the pentane lamp. What legal power they had to do this is not clear, but it might be held to derive from their position under the Act of 1860. The value of the unit was unchanged, at any rate to the order of accuracy with which it had been maintained.

Now that there is no test for the illuminating power of gas in the metropolitan area, the last vestige of legal status for the standard sperm candle or, in fact, for the pentane lamp, seems to have disappeared.

The unsatisfactory nature of the flame standards used by the different countries, viz. the pentane lamp in Great Britain and the U.S.A., the Carcel lamp in France and the Hefner lamp in Germany, led to proposals to use electric lamps as custodians of the unit of candle-power (Fleming 1902). A large number of comparisons of carbon-filament lamps were carried out in the various national laboratories (Paterson 1909), and as a result it was found that the unit employed in the U.S.A. was 1·6 % greater than the British unit, while the French 'bougie décimale' was 1 % greater. Both these countries agreed to adjust their units to conform with the British unit, and in 1909 the three laboratories concerned issued an identical statement in their respective countries to the effect that from 1 April 1909 they would all use the same unit. The text of this statement is given in a paper by Paterson (1910).

The constancy of the unit was assured by repeated exchanges of specially constructed electric lamps between the national laboratories so that the position was fairly satisfactory, at any rate from the practical point of view, although there was no primary standard. In 1921 the name 'International Candle' was given to this agreed unit. Germany, which had not been invited to co-operate, did not come into the scheme and continued to use the Hefnerkerze, which was approximately nine-tenths of a candle.

The present position

The definition of the international candle is incomplete, in that it does not specify how the unit is to be realized in light of any colour other than that of the electric lamps referred to in the definition. These lamps run at a colour temperature slightly below 2000° K, whereas a commercial gas-filled tungsten lamp may run at a colour temperature of 2800° K or more. The completion of the definition had been accomplished implicitly by the Commission Internationale de l'Eclairage in 1924 when it adopted a table of values of the visibility function (the relative sensitivity of the eye to radiant energy of different wave-lengths in the visible spectrum). For some time, however, no attempt was made to use this as a basis for heterochromatic photometry, with the result that the practical units at higher colour temperatures, derived from the carbon-filament lamps at the various national laboratories, showed disagreements of various amounts, the most serious being a progressive change in the ratio of the international candle to the Hefnerkerze from 1·11 to 1·17 on stepping up to a colour temperature of 2600° K.

After a preliminary discussion in 1927, the four national laboratories began co-operative work on the problem of devising a practical means by which the unit at higher colour temperatures could be derived in accordance with the agreed visibility function. As a result of this work it was found that the unit in use in Great Britain, U.S.A. and France for measuring tungsten filament vacuum lamps (i.e. at the colour temperature 2360° K) was too small by about $1\frac{1}{2}$ % and that at higher colour temperatures it was smaller still.

While this work was going on, important developments were taking place, at first at the Bureau of Standards in America and later at the other national laboratories, towards the establishment of a primary standard of light, i.e. a standard quite independent of electric lamps and one reproducible from a written specification.*

This standard consists essentially of a small tube of thoria immersed in a crucible of pure platinum. The tube forms a 'cavity', or 'black-body' radiator, the brightness of which depends solely on its temperature. The temperature of the whole crucible is first raised until the platinum is melted and it is then slowly cooled. The brightness of the mouth of the tube is measured and the value obtained at the 'arrest point', which marks the stage at which the platinum is solidifying, is used to define the unit of luminous intensity. It is, perhaps, fortunate that this standard has a colour temperature, 2046° K, which is close to that of the carbon-filament lamps used to maintain the international candle. The average value found at the various laboratories for the brightness of this standard is 58·9 ± 0·1 international candles per sq.cm.

Until 1929 photometry was not among the subjects dealt with by the International Committee of Weights and Measures, but in that year it was added to the terms of reference of the Consultative Committee on Electricity, and in 1935 a Consultative

---

* *P.V. Com. Int. Poids Mes.* 1931, **14**, 249; 1933, **16**, 254, 256, 261; 1937, **18**, 247.

Committee on Photometry was formed. This Committee recommended the adoption of a new unit of intensity, such that the above black-body standard should be exactly 60 units instead of 58·9. Though this sensibly alters the candle at this temperature the change is not serious for the temperature 2360° K, since the alteration almost exactly compensates the error at that temperature previously referred to.

The international adoption of the new unit was prevented by the war, though it came into force independently in Germany. It is expected that as soon as international co-operation is restored it will become universal.

The other photometric units, viz. the lumen (luminous flux), the foot-candle or metre-candle (illumination; flux received per unit area) and the foot-lambert or apostilb (brightness; flux emitted per unit area of perfect diffuser) are derived by formal definition from the unit of luminous intensity. The only one of these for which material working standards are required is the lumen. The most convenient method of preparing such standards is first to make point-to-point measurements of luminous intensity for a source which is, as nearly as possible, symmetrical about one axis and which can be rotated rapidly about that axis. A tungsten filament vacuum lamp of suitable design is used. The total luminous flux from the lamp can be found by integration from the mean polar curve of luminous intensity thus obtained. Other working standards of luminous flux can then be prepared by means of a spherical photometric integrator.

Standards of radiation

Radiation covering the ultra-violet, visible and infra-red regions of the spectrum can be standardized on a common basis, namely, that of energy. It is convenient to use the electrical units of energy for this purpose, so that the standards of radiation are ultimately based on the units of voltage and resistance.

The evaluation of radiant flux density in absolute units can be obtained by the use of a variety of radiometers. In all of the precision instruments the underlying principle involved is that the radiant energy is degraded into heat by absorption by lamp black or a black-body cavity, and the resulting heat measured either by comparison with heat produced electrically or by the absorption of the heat by Peltier cooling. The efficiency and accuracy of the instruments depend on the closeness with which the conditions of radiative and electrical heating approximate to each other as regards heat loss and distribution.

The accuracy to which the radiant flux density can be determined is very much lower than that attained in the establishment of electrical standards; consequently, the small errors associated with the latter are of no consequence. Scales of radiation maintained at the National Bureau of Standards, the Smithsonian Institute and Uppsala have been compared with the scale maintained at the N.P.L. (Guild 1937) and the general measure of disagreement among these scales probably does not exceed 0·5 %. The internal consistency in the maintenance of the N.P.L. scale is of the order of 0·2 %.

The absolute radiometer, like the majority of absolute instruments, is not convenient as a working instrument and is consequently replaced by thermopiles or calibrated lamps for substandards.

For radiation measurement generally the achievement of an accuracy of even 1 % demands much careful work so that a better accuracy than this is rarely required except by standardizing laboratories.

Spectral energy distribution

In addition to measurements of total radiation it is important to be able to determine the way in which the energy is distributed in the spectrum. This is particularly important in the visible region. When this distribution is known the three response functions, which are internationally agreed as a definition of the normal observer, enable brightness and colour to be derived. Unfortunately, in the visible spectrum the accuracy of direct spectral energy determinations is very much lower than that attained in related measurements such as spectral transmission or reflexion, and is also lower than the discriminating power of the eye. To ensure internal consistency up to visual standards spectral distributions are obtained indirectly by means of tungsten lamps used as working standards of energy distribution. They are calibrated on a colour temperature scale based on visual comparisons with a black-body furnace.

ACOUSTICAL STANDARDS OF MEASUREMENT

BY R. S. DADSON

Introduction

The last few decades have seen rapid advances in the field of applied acoustics which are of profound significance to the community, but which need no enumeration here. In these advances, the increasing perfection of acoustical measuring instruments, and developments in the absolute measurement of sound, have played an important part and are still finding new applications. The present paper attempts to summarize the principles of those absolute methods of measurement which form the basis of the modern science of acoustics, with some discussion of the practical techniques adopted to meet different requirements and the degree of precision now attainable. No attempt is made to treat of the details of individual methods, in regard to which full references are given. Except where otherwise stated, the discussion is restricted to the measurement of sounds of simple-harmonic type of audible frequency, in air under normal atmospheric conditions.

The measurable quantities which together determine the oscillatory conditions at a point in a sound field are:

(*a*) the sound pressure, i.e. the alternating component of the total pressure at the point;

(*b*) the three components of the displacement, or of the particle velocity, relative to the co-ordinate axes.

In practice, however, it is seldom necessary for all of the four quantities referred to in (*a*) and (*b*) to be specified. Largely owing to the fact that the response of the ear is most readily interpreted in terms of sound pressure, the latter is the quantity of greatest significance in practical problems. Measurements of sound pressure are commonly carried out by means of microphones, and the calibration of such microphones in absolute terms is one of the most important applications of the basic measurements. Reference will be made in the section on 'Calibration of microphones' to the methods used at the National Physical Laboratory for this purpose.

Determinations of the acoustical pressure in sound fields are seldom direct; generally they are derived indirectly by calculation from measurements of particle velocity or displacement. In order to obtain satisfactory results by this procedure it is necessary to work with sound fields in which the theoretical relations between pressure and particle velocity are of reasonably simple form and capable of accurate realization in the laboratory. The conditions commonly employed are plane stationary waves in tubes, and progressive spherical waves in an absorbent-lined cabinet. These two procedures are appropriate for the calibration of standard microphones in terms of their 'pressure sensitivity' and 'field sensitivity' respectively, as will be further discussed in the section dealing with this work. By the use of devices such as the pistonphone it is possible to generate known acoustical pressures in a small enclosure over the lower part of the audio-frequency range, and thence to obtain absolute pressure calibrations of suitable microphones. This method is independent of the above-mentioned indirect methods, and since the ranges of frequency over which the two procedures are applicable overlap, a useful check of accuracy is made possible.

It is not within the scope of this paper to give details of the different varieties of microphones in use for practical sound measurements. They may be of the condenser, moving-coil, piezo-electric, or pressure-gradient types, all of which have their special applications. Of these the condenser microphone, introduced by Wente (1917), has been found to be particularly suitable for use in the laboratory for standardization purposes. This it owes to its simple and comparatively rigid construction, the possession of a flat and extremely stiff diaphragm which, for many practical purposes, may be regarded as a rigid boundary, its reasonably high sensitivity, and above all its great stability of calibration over long periods of time. There has lately been an increasing tendency to use piezoelectric microphones also as standards. Thus in practice a condenser microphone carefully calibrated in absolute terms may be regarded virtually as having the status of an absolute standard for the measurement of sound pressures under the appropriate experimental conditions. The unit of sound pressure is the dyne/cm.^{2}.

Absolute measurement of particle velocity: the Rayleigh disk

For the purpose of the absolute measurement of particle velocity at a point in a sound field the Rayleigh disk is widely used. The development of accurate acoustical measurements has been very closely bound up with the study of this most useful and simple device. Its action is based on the well-known fact that a flat object suspended in a stream of fluid is acted upon by a couple which tends to turn it at right angles to the direction of flow. It was proposed by Lord Rayleigh (1880, 1882) that this effect should be turned to account for the absolute measurement of sound, and Koenig (1891) obtained the following expression for the torque T on an infinitely thin circular disk, which has formed the basis of subsequent work:

$$T = \tfrac{1}{6}\rho d^3 v^2 \sin 2\theta. \tag{1}$$

In this expression d, the diameter of the disk, is assumed to be small compared with the wave-length of the sound, and the density of the material of the disk is assumed to be very large compared with ρ, the density of the medium. v is the root-mean-square value of the particle velocity in the sound wave, and θ the angle between the normal to the disk and the direction of flow. The medium is assumed to be frictionless and the flow laminar.

The disk, carefully shielded from draughts, is suspended in the sound field on a fine fibre usually of glass or quartz, measurements being made either of the torque required to maintain the disk at a fixed inclination to the lines of flow, or of the actual deflexion of the disk in response to the sound. The value of θ is normally chosen to be 45°, at which value the sensitivity of the device is a maximum. Practical details of the methods of calibrating and using Rayleigh disks will be found in a particularly useful paper by Barnes & West (1927).

In the application of the Rayleigh disk a number of small departures from Koenig's formula are encountered which have been studied in detail by many workers. A very comprehensive investigation has been made by Scott (1945). Zernov (1908) and Scott have shown experimentally that on account of the finite thickness of the disk the torque tends to exceed the value predicted by Koenig's formula to an extent which increases with the ratio of thickness to diameter. According to Scott the excess varies from about 1 to 10 % in torque (equivalent to 0·5–5 % in particle velocity) over a range of the ratio of thickness to diameter of from 0·006 to 0·12. To minimize this effect, therefore, the disks should be as thin as possible. A small tendency in the opposite direction is also found which originates from the fact that, owing to lack of infinite inertia, the disk is not completely immobile under the influence of the fluctuating torque. Appropriate corrections have been derived by Wood (1935) and King (1935), but, for most practicable types of disks used in air, would only amount to the order of 1 % in torque (0·5 % in particle velocity). King has also given a comprehensive theoretical treatment of the effect of the diffraction of the sound by the disk. The departure from Koenig's formula so arising, which is only appreciable at the higher frequencies, reaches about 1 % excess in torque (0·5 % in particle velocity) when the wave-length falls to some 15 times the diameter of the disk, increasing approximately as the square of the frequency. In the case of disks used in a plane-wave system in a tube, there is a further small effect due to the proximity of the walls, which, according to Scott's measurements, may amount to an excess of some 2 % in torque (1 % in particle velocity) when the diameter of the disk reaches about one-third of that of the tube. Barnes & West have shown that appreciable errors may arise over narrow bands of frequency close to the internal resonances of the disk; these errors, however, are very localized in frequency and may be eliminated if care is taken to avoid using any particular disk at frequencies close to these resonances.

Barnes & West have concluded that if normal precautions are taken to ensure that the conditions presupposed in Koenig's formula may reasonably be expected to be valid, the measurement of particle velocity may be effected with the Rayleigh disk to an accuracy within 2 %.

Merrington & Oatley (1939) and Scott have suggested as a result of their investigations that there is a residual uncertainty in the constant in Koenig's formula which may lead to over-estimation of particle velocity of up to about 3% over certain ranges of frequency. The origin of this effect is so far obscure and appears to deserve further study. Nevertheless, this conclusion is not greatly at variance with the figure of 2 % estimated by Barnes & West. These estimates of the order of accuracy of Rayleigh disk measurements are in line with the experience over a period of years at the N.P.L. in the use of the instrument as a reference standard for the calibration of microphones. In the use of a disk for the comparison of sounds of identical frequency it is possible that the accuracy may be somewhat improved, in that uncertainties in the constant in Koenig's formula will largely be eliminated.

During the history of its development and use the Rayleigh disk has proved itself within the limits mentioned to be a reliable instrument, the performance of which justifies its continued use as a reference standard.

Absolute measurement of displacement: the smoke particle method

Carrière (1929), Andrade (1931), and subsequently Andrade & Parker (1937), have investigated the use of illuminated smoke particles for the measurement, by direct observation, of the displacement amplitude of sound waves in a tube over a wide range of frequency. Theoretical considerations indicate that particles of sufficiently small cross-section and inertia introduced into the sound field may be regarded as vibrating with very nearly the full amplitude of the wave motion. In practice the observations are made in a plane-wave system in a tube. The particles are illuminated by a source such as an arc lamp and their excursions observed by microscope or by photographic means. Of various substances tried, magnesium oxide smoke has proved especially suitable.

Scott employed this method as a reference standard in his investigation of the residual errors of the Rayleigh disk, and it has also been used by Cook (1940) for assessing the accuracy of the thermophone and reciprocity methods for the absolute pressure calibrations of microphones, which will be discussed in a later section.

As regards precision, it appears that an accuracy of 1 % or better, in the measurement of displacement amplitude, may be realized over the range of frequency for which the motion of the smoke particles may reasonably be regarded as indistinguishable from that of the medium. For particles of magnesium oxide smoke of radius about $0\cdot3\mu$, which is typical of the sizes found in practice, this condition would be expected to hold up to a frequency of about 4000 cyc./sec., with an error of perhaps 1 or 2 % at 5000 cyc./sec. (Scott). Since, for given sound pressure or particle velocity, the displacement amplitude is inversely proportional to the frequency of the sound, it becomes necessary to use somewhat intense sounds at the upper frequencies in order to obtain sufficiently long traces for measurements of the highest accuracy. It is also worthy of note that the smoke method as at present used yields the value of the total excursion (peak to peak) of the particles. Purity of wave form is therefore of particular importance if it is desired to relate the results to quantities such as the root-mean-square pressure or particle velocity determined by other methods.

The technique of observation with smoke particles, although of comparatively recent introduction, has already proved of great value in sound measurement, and gives as close an approach to a truly absolute measurement as any method yet available.

Absolute determination of sound pressure and allied measurements

In so far as the techniques described in the sections on absolute measurement of particle velocity and displacement relate to measurement of pressure, they have relied on the absolute determination of the particle velocity or displacement in sound waves of simple form (such as stationary plane waves or progressive spherical waves) and the calculation therefrom of the sound pressure making use of the known relations between the various quantities. Consideration is now given to an independent class of absolute measurements based on the generation of known acoustic pressures in a small rigid-walled enclosure, by means of which it is possible to obtain directly the pressure calibrations of microphones of certain types. Two pressure generators of this nature, known respectively as the thermophone and pistonphone, have been widely used for the calibration of condenser microphones; it is convenient to consider these two instruments together.

The satisfactory operation of this method depends on obtaining a high degree of uniformity of sound pressure throughout the enclosure. The range of application is thus restricted to frequencies at which the wave-length is large compared with the dimensions of the enclosure.

The thermophone, introduced by Arnold & Crandall (1917) and further developed by Wente (1922), Ballantine (1932) and others, consists of a small metal-walled enclosure containing strips of gold or platinum foil through which is passed a.c. superposed on a much larger direct current. The resultant heating and cooling of the strip generates an alternating pressure in the enclosure which can be calculated provided the many constants of the instrument are accurately known and correct assumptions regarding the *modus operandi* of the system are made. In practice, however, these calculations are very complicated, and some of the constants and correction factors involved are not altogether free from uncertainty.

The pistonphone operates on much simpler principles. The alternating pressure in the enclosure is generated by the simple-harmonic vibration of a piston, the amplitude of which is either prescribed by mechanical means or directly measured. The calculation of the pressure is comparatively straightforward, and the physical constants involved are few in number and easy to determine.

Several authors (Wente, Ballantine, Cook, Ernsthausen and others) have assessed the performance of the thermophone and pistonphone by using them to calibrate condenser microphones which have also been calibrated by independent means. The pistonphone has generally proved a reliable instrument, and there appears to be general agreement that it may be relied upon to give an accuracy within 3 or 4 %. On the other hand, published accounts of the use of the thermophone are inconsistent as to its precision. Whilst certain workers (e.g. Wente, Ballantine) have succeeded in using the thermophone as an absolute instrument with a precision comparable with that of the pistonphone, others (e.g. Cook) have reported considerable discrepancies which they attribute to uncertainties in the behaviour of the thermophones they employed.

Two other methods, which will only be mentioned briefly, may conveniently be grouped with the pistonphone and thermophone; both these methods allow of the calibration of microphones of suitable types by means of purely electrical measurements. The first, introduced by Gerlach (1923) and further developed by Ballantine, is often called the electrostatic actuator method, and was designed for the calibration of condenser microphones. An additional electrode in the shape of a grille is placed close to the diaphragm and an alternating e.m.f. superimposed on a large, steady polarizing e.m.f. is applied between the two. This results in an alternating force on the diaphragm which may be calculated from the form and dimensions of the electrode. This method is reputed to be comparable in accuracy with the pistonphone. The second method, which has received increasing attention during recent years, is based on the principle of reciprocity as applied to linear reversible electro-acoustic transducers.* Accounts of the method have been published by Maclean (1940), Foldy & Primakoff (1945), Cook and others; the latter author has made an extensive investigation of the application of reciprocity to the calibration of microphones of the condenser and crystal types, and has found the results to agree generally with those afforded by the pistonphone and electrostatic actuator to an accuracy of the order 4 or 5 %. The principle of reciprocity appears to offer considerable promise as a supplement to the existing methods.

Calibration of microphones

The absolute measurements discussed in the foregoing sections are suitable only for fundamental laboratory determinations, and most practical sound measurements are carried out by means of microphones, associated with electrical amplifying and indicating equipment. The calibration of such microphones is therefore a

---

* In its simplest form the method is analogous to the well-known Hopkinson test of the performance of electrical machines.

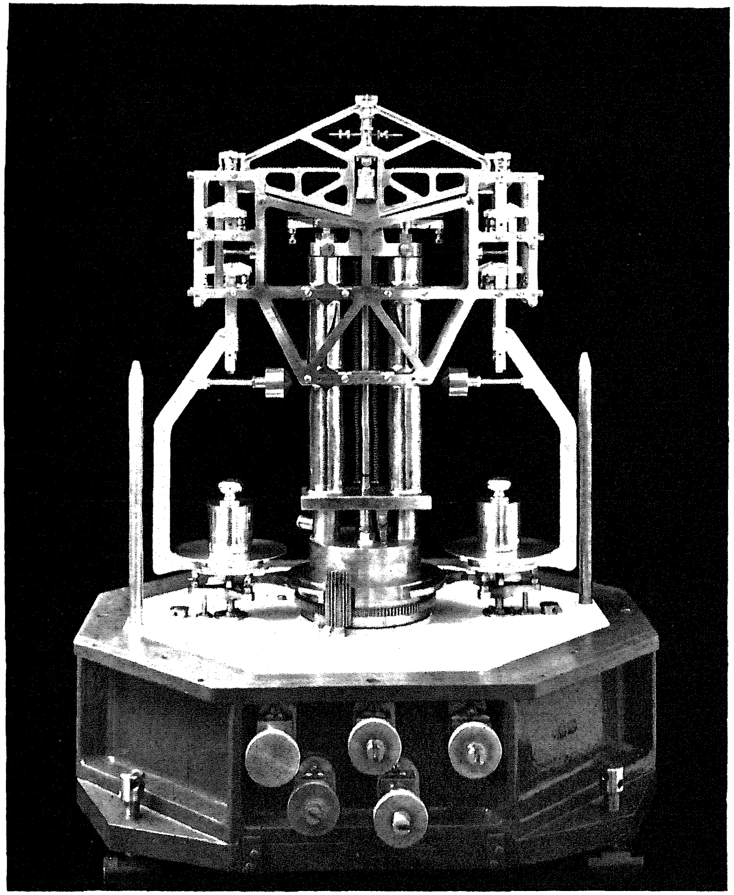


FIGURE 1. Precision balance at The National Physical Laboratory. Cover removed.

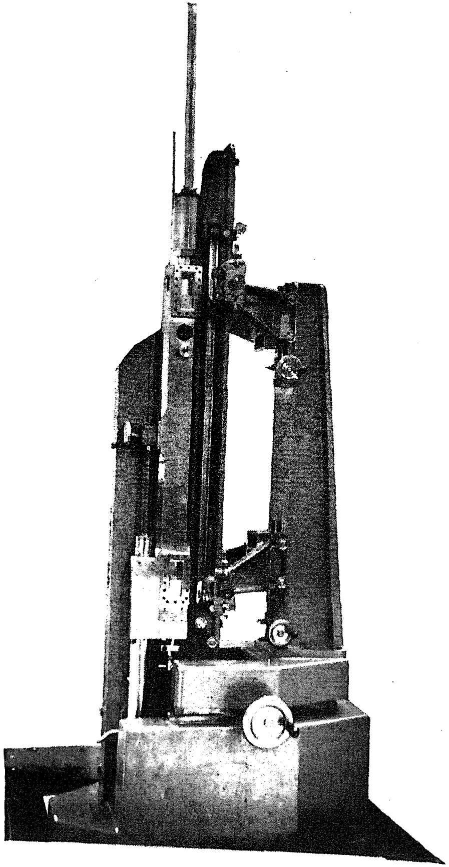


FIGURE 5. National Physical Laboratory primary standard barometer.

particularly important application of the basic measurements. Even for purposes of microphone calibration it is not always necessary to refer directly to absolute methods, and it has been the practice at the N.P.L. to maintain certain high-quality microphones, carefully calibrated in absolute terms, as reference standards against which other microphones may be compared. This procedure gives ample accuracy for ordinary purposes as well as resulting in simplification and saving of time. The microphones chosen as standards are of the condenser type (Wente 1917; Oliver 1930), which have been found by repeated tests to maintain stability of calibration over very long periods.

Owing to the disturbance of the sound field caused by the presence of a microphone, including the resonance of the cavity normally existing in front of the diaphragm, it is necessary to distinguish between two different ways of expressing the sensitivity (Aldridge 1928; West 1929). There is, first, its 'field sensitivity', which is the ratio of the response of the microphone when placed in a progressive sound wave to the sound pressure which existed at the same position before the introduction of the microphone. This quantity determines the performance of the microphone when employed to investigate free sound fields, and is the calibration appropriate to most ordinary sound measurements. Secondly, there is the 'pressure sensitivity' which is the ratio of the response of the microphone to the actual pressure actuating its diaphragm. This form of calibration determines the performance of the microphone when used to determine the pressure set up in a small enclosure, as, for instance, in the case of an artificial ear for the measurement of the acoustical output of telephone receivers.

The determination of the field sensitivity of the standard microphone is carried out, between the frequencies 300 and 10,000 cyc./sec., in a progressive wave in a cabinet heavily lagged with sound-absorbing material, in one side of which is situated a loud-speaker serving as the source of sound. Measurements of the sound particle velocity are made by means of the Rayleigh disk at a series of points on the axis of the loud-speaker and are converted into the corresponding sound pressures from the known relations in the progressive wave. The microphone is substituted for the disk and its response measured at the same positions. At frequencies below 300 cyc./sec. reflexions from the walls of the cabinet are no longer negligible, with the result that the method ceases to be valid. However, at these low frequencies the dimensions of the standard microphone are small compared with the wavelength of the sound, and since the field and pressure sensitivities then become practically indistinguishable, the determination of the pressure sensitivity suffices to determine the field sensitivity also. The Rayleigh disks employed at the N.P.L. are normally of glass, 1 cm. in diameter and of the order of 0·015 cm. thick, and are suspended on quartz fibres about 5μ in diameter. Smaller disks, 0·5 cm. in diameter, are also employed at the upper frequencies. For the accurate determination of frequency, reference is made to the N.P.L. Primary Frequency Standard (Essen 1938.)

The pressure sensitivity is determined from Rayleigh disk measurements in a stationary wave system set up in a cylindrical tube, the diameter of which is usually chosen to be approximately equal to that of the microphone diaphragm, which closes one end of the tube. The stationary waves are set up by a pure-tone source at the other end of the tube, the Rayleigh disk being situated at an odd number of quarter wave-lengths from the microphone diaphragm, i.e. very closely at a velocity antinode. From the particle velocity determined by the deflexions of the disk the sound pressure actuating the diaphragm of the microphone may be deduced from the known relations in the stationary wave, and compared with the microphone response. Care must be taken in practice to avoid transverse resonances in the tube; in the case of the particular design of microphone concerned, the diaphragm of which is about 5 cm. in diameter, such resonances may be encountered in the frequency range above about 3500 cyc./sec. The use of stationary waves in hydrogen, in which the wave-length at a given frequency is nearly four times its value in air, may provide a means of extending the range of such measurements to higher frequencies.

The stationary wave method is not employed at the N.P.L. at frequencies below 62·5 cyc./sec. on account of the inconveniently long tube required, and at the lower end of the frequency range, from 10 to 400 cyc./sec., the pressure calibrations of the standard microphones have been determined by means of a moving-coil pistonphone.

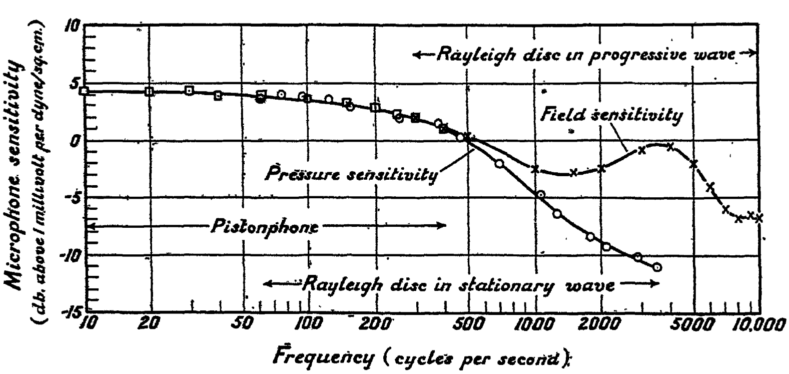


FIGURE 6. Calibration curves of standard condenser microphone.

Calibration curves determined by Mr N. Fleming for one of the standard condenser microphones in use at the Laboratory are shown in figure 6. Over the range 62·5–400 cyc./sec. the pressure sensitivity has been determined both with the Rayleigh disk and pistonphone, the agreement between the results of the two methods being of the order of 3 % which may be considered satisfactory. Good agreement is also shown between the pressure and field sensitivities, determined by the Rayleigh disk in stationary and progressive waves respectively, in the region 300–400 cyc./sec., where the two curves tend to become indistinguishable.

The applications of absolute measurements: measurement of noise

The application to industrial and other practical problems of the absolute measurements discussed in the preceding sections virtually reduces to the use of microphones which have been calibrated in absolute terms, usually by comparison with microphones maintained as laboratory standards. These applications vary so widely in nature that the question of the precision of the measurements involved can only be treated in the broadest possible way. Frequently the sounds concerned are very complex, and the circumstances in which they arise can only be approximately specified. An accuracy closer than about 10 %, that is to say of the order 1 db.,* is rarely demanded in such measurements.

In recent years much attention has been paid to the problems of measurements associated with the suppression of noise, and in 1936 the British Standards Institution laid down a standard procedure for the measurement of noise by subjective means, and defined the terms to be adopted. In brief, the loudness of noise is judged by comparison, in free air, with a standard pure tone of frequency 1000 cyc./sec. under certain specified experimental conditions. If the 'normal' observer (i.e. an average of a group of persons of normal hearing) decides that a noise is equally loud to a 1000 cyc./sec. tone of intensity n db. above a reference level corresponding to a root-mean-square sound pressure of 0·0002 dyne/sq.cm. (which is very close to the normal threshold of hearing for the standard frequency), the noise is said to have an 'equivalent loudness' of n British Standard phons. An International Commission held in Paris in 1937 reached provisional agreement on the international standardization of the phon scale. It would be outside the scope of this paper to discuss the question of the standard procedure for such noise measurements in any detail, in which connexion the British *Standard Glossary of Acoustical Terms and Definitions* (1936) and a paper by Churcher & King (1937) should be consulted. The need for suitable objective instruments for the measurement of noise has also long been realized, and useful progress has been made in the design of objective meters employing a microphone, amplifier and suitable networks simulating the characteristics of the ear, and designed to give readings in phons for various types of noises of common occurrence (Davis 1938; King, Guelke, Maguire & Scott 1941). The precision with which such measurements may, in general, be made and interpreted is difficult to assess owing to the widely differing types of noise encountered, and the variety of circumstances, psychological and otherwise, in which they occur. In the more restricted case of reasonably steady noises, an accuracy of the order 2 to 3 phons should usually be possible if suitable precautions are taken.

The author is indebted to Mr N. Fleming for information concerning some of the past work on acoustical standards at the National Physical Laboratory.

---

* The decibel notation is commonly employed in acoustics on account of the very wide range of intensity which can be accommodated by the ear. Two powers W_1 and W_0 are said to be separated by an interval of n db. where $n = 10 \log_{10} (W_1/W_0)$. If the conditions are such that the ratio of the pressures P_1 and P_0 (or of the corresponding particle velocities) is the square root of the power ratio, then $n = 20 \log_{10} P_1/P_0$ also. In most practical circumstances a variation of less than 1 db. is scarcely noticeable to the ear.

REFERENCES

Aldridge 1928 *J. P.O. Elect. Engrs,* **21**, 223.

Andrade 1931 *Proc. Roy. Soc.* A, **134**, 445.

Andrade & Parker 1937 *Proc. Roy. Soc.* A, **159**, 507.

Arnold & Crandall 1917 *Phys. Rev.* **10**, 22.

Ballantine 1932 *J. Acoust. Soc. Amer.* **3**, 319.

Barinov 1941 Monograph in Russian. *The present state of standards of length and methods of accurate measurement of length.* Leningrad: All-Union Scientific Research Institute of Metrology.

Barrell 1945 *Machine Shop Mag.* **6**, no. 12, 36.

Barrell & Sears 1939 *Phil. Trans.* A, **238**, 1.

Barnes & West 1927 *J. Instn Elect. Engrs,* **65**, 871.

Beattie 1941 *Temperature, its measurement and control in science and industry.* New York: Amer. Inst. Phys.

Beattie & Blaisdell 1935 *Proc. Amer. Acad. Arts Sci.* **71**, 360.

Benoît & Chaney 1895 *Trav. Bur. int. Poids Mes.* **12**, 24.

Benoît, Fabry & Perot 1913 *Trav. Bur. int. Poids Mes.* **15**, 3.

Birge 1941 *Rep. Progr. Phys.* **8**, 131.

Blaisdell & Kaye 1941 *Temperature, its measurement and control in science and industry.* New York: Amer. Inst. Phys.

B.S.I. 1936 *Standard glossary of acoustical terms and definitions,* B.S.S. 661.

Browne 1936 *Spec. Publ. U.S. Cst Geod. Surv.* no. 204.

Browne & Bullard 1940 *Proc. Roy. Soc.* A, **175**, 110.

Bullard & Jolly 1936 *Mon. Not. R. Astr. Soc.* (Geophys. Suppl.), **3**, 443.

Bullard & Jolly 1937 *Mon. Not. R. Astr. Soc.* (Geophys. Suppl.), **4**, 132.

Callendar 1899 *Phil. Mag.* **47**, 191.

Carrière 1929 *J. Phys. Radium,* **10**, 198.

Chappuis 1888 *Trav. Bur. int. Poids Mes.* **6**.

Churcher & King 1937 *J. Instn Elect. Engrs,* **81**, 57.

Clarke 1867 *Phil. Trans.* **157**, 161.

Clarke 1873 *Phil. Trans.* **163**, 445.

Clark 1939 *Phil. Trans.* A, **238**, 65.

Cook 1940 *Bur. Stand. J. Res.* **25**, 489.

Curtis, H. L. 1944 *Bur. Stand. J. Res.* **33**, 235.

Davis 1938 *J. Instn Elect. Engrs,* **83**, 249.

Day & Sosman 1911 *High temperature gas thermometry.* Washington: Carnegie Inst.

Dorsey 1940 *The properties of ordinary water substance.* New York: Reingold.

Dryden 1942 *Bur. Stand. J. Res.* **29**, 303.

Ernsthausen 1939 *Akust. Z.* **4**, 13.

Essen 1938 *Proc. Phys. Soc.* **50**, 413.

Fleming 1902 *J. Instn Elect. Engrs,* **32**, 119.

Foldy & Primakoff 1945 *J. Acoust. Soc. Amer.* **17**, 109.

Gerlach 1923 *Wiss. Veroff. Siemens-Konzern,* **3**, 139.

Guild 1937 *Proc. Roy. Soc.* A, **161**, 1.

Guillaume 1927 *La Création du Bureau International et son Œuvre.* Paris: Gauthier-Villars.

Hall 1930 *Phil. Trans.* A, **229**, 1.

Heycock & Neville 1895 *J. Chem. Soc.* **67**, 160.

Heyl & Cook 1936 *Bur. Stand. J. Res.* **17**, 805.

Holborn & Day 1900 *Ann. Phys., Lpz.,* **2**, 505.

Kater 1818 *Phil. Trans.* p. 103.

Kaye 1911 *Proc. Roy. Soc.* A, **85**, 430.

Kaye 1936 *J. Acoust. Soc. Amer.* **7**, 167.

King, A. J., Guelke, Maguire & Scott 1941 *J. Instn Elect. Engrs,* **88**, 163.

King, L. V. 1935 *Proc. Roy. Soc.* A, **153**, 17.

Koenig 1891 *Ann. Phys. Chem.* **43**, 43.
Kösters 1938 *Werkstattstechnik*, **32**, 527.
Kösters & Lampe 1934 *Phys. Z.* **35**, 233.
Kösters & Sears 1935 *P.V. Com. int. Poids Mes.* **17**, 113.
Kühnen & Furtwängler 1906 *Bestimmung der Absoluten Grösze der Schwerkraft zu Potsdam.* Potsdam: Geod. Inst.
Maclean 1940 *J. Acoust. Soc. Amer.* **12**, 140.
Mallett & Dutton 1925 *J. Instn Elect. Engrs*, **63**, 502.
Manley 1933 *Phil. Mag.* **16**, 489.
Manley 1935 *Phil. Mag.* **19**, 243.
Manley 1945 *Proc. Phys. Soc.* **57**, 136.
Meggers & Peters 1918–19 *Bull. U.S. Bur. Stand.* **14**, 697.
Merrington & Oatley 1939 *Proc. Roy. Soc.* A, **171**, 505.
Michelson & Benoît 1895 *Trav. Bur. int. Poids Mes.* **11**, 3.
Moser 1930 *Ann. Phys., Lpz.*, **6**, 852.
N.P.L. Annual Reports 1934, 1935.
Oliver 1930 *J. Sci. Instrum.* **7**, 113.
Paterson 1909 *Phil. Mag.* **18**, 263.
Paterson 1910 *Coll. Res. Nat. Phys. Lab., Lond.*, **6**, 115.
Pérard 1934 *Trav. Bur. int. Poids Mes.* **19**, 66.
Poynting 1878 *Proc. Roy. Soc.* **28**, 2.
Rayleigh 1880 *Proc. Camb. Phil. Soc.* **4**, 18.
Rayleigh 1882 *Phil. Mag.* **14**, 186.
Rolt & Barrell 1927 *Proc. Roy. Soc.* A, **116**, 417.
Rolt & Barrell 1929 *Proc. Roy. Soc.* A, **122**, 122.
Scott 1945 *Proc. Roy. Soc.* A, **183**, 296.
Sears 1936 *Sci. Progr.* no. 122, 209.
Sears & Barrell 1932 *Phil. Trans.* A, **231**, 75.
Sears & Barrell 1934 *Phil. Trans.* A, **233**, 143.
Sears & Clark 1933 *Proc. Roy. Soc.* A, **139**, 130.
Sears, Johnson & Jolly 1928 *Phil. Trans.* A, **227**, 281.
Smith 1922 Article in *Dictionary of Applied Physics*, ed. Glazebrook, **2**, 230, 241. London.
Thiesen, Scheel & Dieselhorst 1900 *Wiss. Abh. Phys.-tech. Reichanst., Berl.*, **3**, 68.
Tilton & Taylor 1937 *Bur. Stand. J. Res.* **18**, 205.
Tutton 1931 *Phil. Trans.* A, **230**, 293.
Watanabe & Imaizumi 1928 *Proc. Imp. Acad. Japan*, **4**, 350.
Wente 1917 *Phys. Rev.* **10**, 39.
Wente 1922 *Phys. Rev.* **19**, 498.
West 1929 *J. Instn Elect. Engrs*, **67**, 1137.
Williams 1933 *Proc. Phys. Soc.* **45**, 699.
Wood 1935 *Proc. Phys. Soc.* **47**, 779.
Zernov 1908 *Ann. Phys., Lpz.*, **26**, 79.

Kinetic studies in the chemistry of rubber and related materials

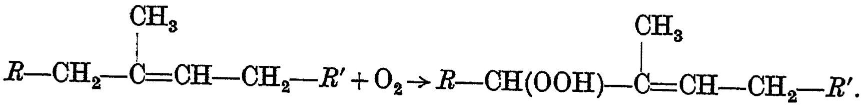
I. The thermal oxidation of ethyl linoleate

By J. L. Bolland

(*Communicated by Eric K. Rideal, F.R.S.—Received* 5 *March* 1945)

The kinetics of the initial stages of the thermal oxidation of ethyl linoleate (by molecular oxygen) have been investigated in the temperature range 35–75° C. From consideration of chemical and kinetic evidence the reaction mechanism has been established: oxidation chains are initiated by thermal decomposition of ethyl linoleate hydroperoxide (which in the early stages of oxidation is substantially the sole product). The chain propagation reactions are identified as $R\text{—} + O_2 \rightarrow RO_2\text{—}$ and $RO_2\text{—} + RH \rightarrow ROOH + R\text{—}$ (where RH represents ethyl linoleate). Chain termination occurs by mutual destruction of the radical chain carriers, R— and RO_2—.

In recent years the chemistry of unconjugated polyolefins has been extensively studied, and it has been abundantly demonstrated that a very important type of reactivity is that associated with methylene groups adjacent to a double bond (Farmer 1942). This α-methylenic reactivity plays an essential role in many of the reactions of natural rubber, including the important one of oxidation: thus when rubber is oxidized photochemically the first identifiable product is a hydroperoxide of undiminished unsaturation (Farmer & Sundralingam 1943), which on the analogy of the characteristic behaviour of a series of simple unconjugated olefins (Criegee, Pilz & Flygare 1939; Farmer & Sundralingam 1942; Farmer & Sutton 1942, 1943; Sutton 1944) may be represented thus:



$$R\text{—}CH_2\text{—}\overset{\displaystyle CH_3}{\overset{|}{C}}\text{=}CH\text{—}CH_2\text{—}R' + O_2 \rightarrow R\text{—}CH(OOH)\text{—}\overset{\displaystyle CH_3}{\overset{|}{C}}\text{=}CH\text{—}CH_2\text{—}R'. \quad (1)$$

While photo-oxidation of such olefins does follow this course almost quantitatively in the initial stages, secondary reactions involving these primary products soon become of importance; in the case of thermal oxidation it is not in general possible to find conditions which make quantitative yields of hydroperoxide possible even at the outset. In embarking on a series of kinetic investigations designed to throw light on the precise mechanism of this type of oxidation, attention has in the first instance been devoted to olefins which can be oxidized thermally at a convenient rate and with quantitative production of hydroperoxide. The 1:4 olefinic esters appear to fulfil this requirement admirably, and the present paper describes an examination of the kinetics of the thermal oxidation of one of the simplest examples of this type, ethyl linoleate (I):

$$CH_3(CH_2)_4.CH\text{=}CH\text{—}CH_2\text{—}CH\text{=}CH\text{—}(CH_2)_7.COOEt. \qquad \text{(I)}$$

It will be shown that the oxidation products in the early stages of oxidation are simple and correspond exclusively to the type of reaction indicated by (1). The possibility thus arises of examining what appears to be the key reaction in the oxidation of many olefins of various types in absence of complicating secondary reactions.

EXPERIMENTAL

Ethyl linoleate was prepared from a sample of cotton-seed oil by the bromination-debromination procedure. The tetrabromostearic acid was repeatedly crystallized from ethyl acetate to give a product of melting-point 115·7 ± 0·1° C and bromine content 53·3 % (theoretical 53·3 %). The ethyl linoleate so obtained was carefully fractionated in a molecular still at 78° C. The final product had a peroxide content < 0·01 %, and as it was stored carefully *in vacuo*, this value did not increase appreciably during the investigation.

Ethyl stearate was prepared by esterifying B.D.H. stearic acid ('purified stearine') and recrystallizing the ester three times from ethanol. The final product melted at 33·3 ± 0·1° C.

Benzoyl peroxide was recrystallized three times from ether, when a sample of peroxide content 6·42 % (theoretical 6·61 %) was obtained.

Rates of oxidation were measured in an apparatus of the constant-pressure type, indicated in figure 1. As the supply of ethyl linoleate available was strictly limited, the apparatus was designed to give accurate measurements of oxidation rates using a sample of about 0·2 g. The gas burette *A* (internal cross-section about 1 sq.cm. and length 20 cm.) and controlling manometers *B* and *C*, were immersed in a water thermostat, operating at 25·00 ± 0·01° C, and were connected to the reaction vessel *D* (*ca.* 1–5 c.c. volume) by means of a 6 in. length of Thiokol tubing of external and internal diameters 15 and 2 mm. respectively. Vigorous agitation of the contents of the reaction vessel was achieved by connecting to an eccentric wheel driven at speeds of about 200 r.p.m. The reaction vessel was immersed in a subsidiary thermostat, the temperature of which was controllable to 0·05° C.

Prior to each run the required amount of ethyl linoleate (normally *ca.* 0·2 g.) was weighed into the nitrogen-filled reaction vessel and the glass-thiokol connexions sealed with picein. The apparatus was then thoroughly evacuated at 10^{-5} mm. and left overnight with free access to a liquid-oxygen trap.

During an oxidation run the di-octyl phthalate manometer *B* was used to indicate deviations from the chosen constant oxygen pressure in the apparatus; by observing the meniscus through a reading microscope (magnification 10 ×), pressure alterations of 0·005 mm. Hg could readily be detected. The upward movement of the mercury level in the gas burette was actuated by the evolution of gas from the small electrolytic cell *E*. The method of automatic control of the gas burette, originally described by Kohman (1929), was used. A thermionic relay, based on the original design of Hershberg & Huntress (1933), was incorporated in the device. Details of the circuit, designed by Dr G. F. Bloomfield of these laboratories,* are included in figure 1. Since the current through the platinum control contacts sealed into the reverse side of the wide-limbed mercury manometer was only a few microamperes, the repeated making and breaking of the contacts, even in an atmosphere of oxygen, caused no sign of contamination of the mercury surface. Provided the current through the electrolytic cell was regulated so that the cell was on and off for approximately equal periods, the pressure variation during the control cycle did not exceed 0·02 mm. During an oxidation run, therefore, the automatic device was employed to maintain the oxygen pressure close to the selected standard value, while at regular intervals the exact time at which the standard pressure was obtained—as indicated on the octoil manometer—was noted and the current burette level measured to 0·01 mm.

---

* The British Rubber Producers' Research Association.

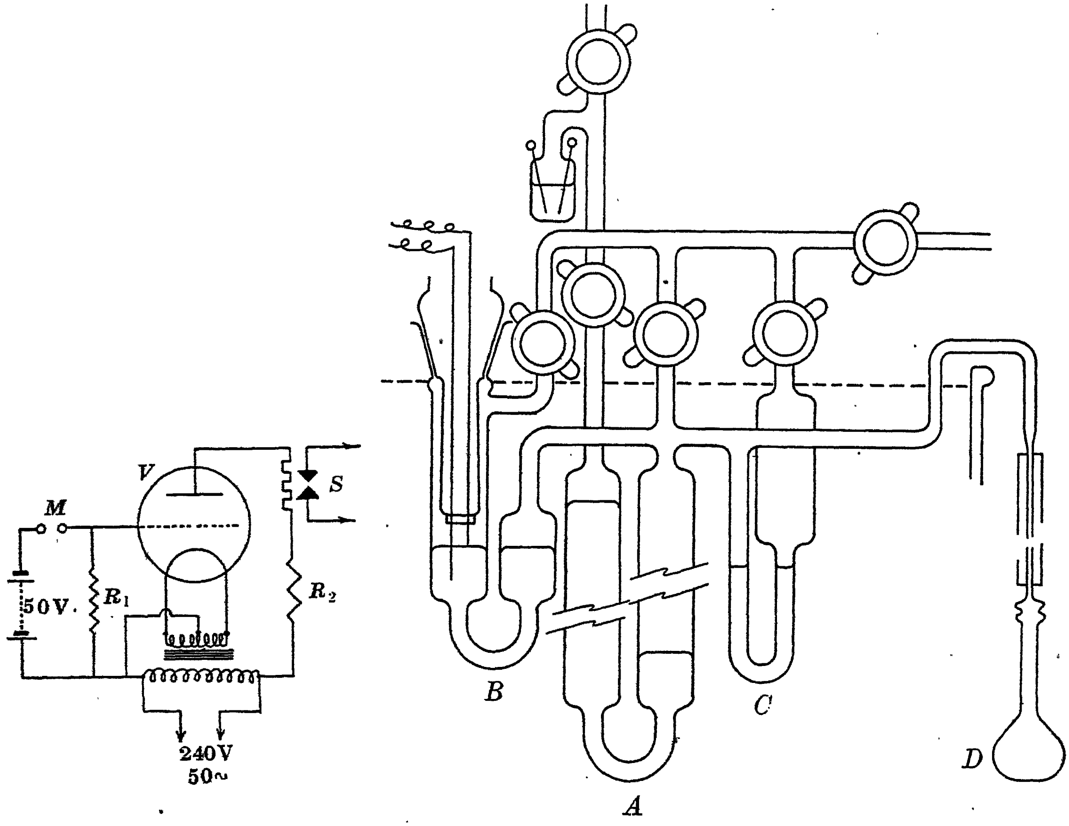


FIGURE 1. Apparatus for measuring rates of oxidation and circuit of electronic relay. V represents a PX4 valve, S a Sunvic hot-wire vacuum switch (type 602V), M the contacts in the manometer, C and R_1 and R_2 resistances of 5 megohms and 5000 ohms respectively.

The thiokol tubing, forming the flexible connexion to the reaction vessel, was baked out *in vacuo* for 16 hr. at 100° C before use; after such treatment only negligible amounts of permanent and condensable gases penetrated into the apparatus over periods of many days.

The peroxide content of oxidation products was, when possible, determined by the colorimetric ferrous thiocyanate method described by Bolland, Sundralingam, Sutton & Tristram (1941) and the iodometric method of Dastur & Lea (1941). Both methods were found capable of giving reproducible results, but the iodometric method suffered the disadvantage of requiring an inconveniently large sample if accurate results were to be obtained for peroxide contents less than 0·1 g.mol. peroxide/g.mol. ethyl linoleate.

Standard experimental conditions, under which the oxidation of ethyl linoleate proceeded at a readily measurable rate, were chosen as follows: temperature, 45° C; oxygen pressure, 100 mm. Hg; concentration of ethyl linoleate, 100 %; and weight of linoleate sample, 0·2 g.

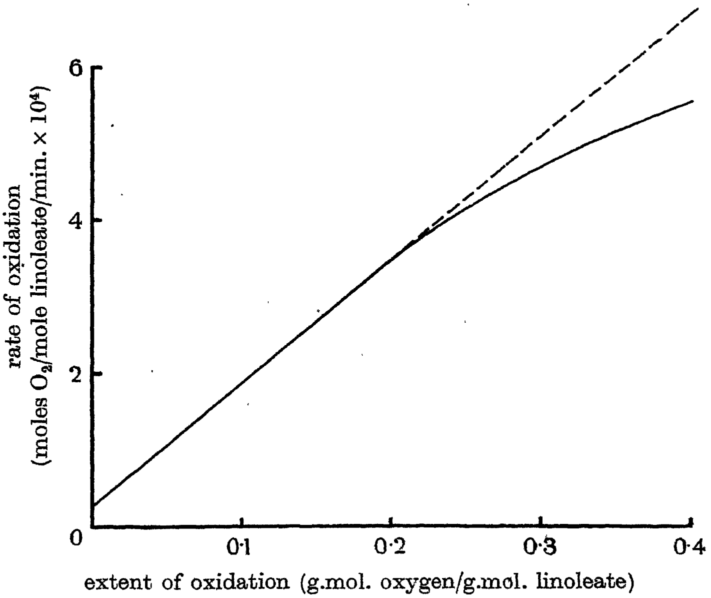


FIGURE 2. Oxidation of ethyl linoleate at 25° C and 100 mm. oxygen pressure.

The course of a typical oxidation run under such conditions is shown in figure 2. Oxidation commences without any apparent inhibition period. During the introduction of up to 0·2 g.mol. oxygen/g.mol. ethyl linoleate a linear relation between the experimentally determined rates of oxidation and the extent of oxidation is satisfied. At higher extents of oxidation the curve deviates from linearity until at about 1·0 g.mol. oxygen/g.mol. linoleate the maximum rate of oxidation is reached. In this investigation, attention has been concentrated on the earlier stages of oxidation, during which, there is reason to believe, the complications associated with secondary reactions—such as peroxide decomposition and molecular intercondensation—are of minor importance.

Formation of volatile products. In the case of a sample of ethyl linoleate oxidized at 45° C the formation of products condensable in liquid oxygen was found to correspond to 1 mol. for every 70 and 100 mol. oxygen reacted during the introduction of the first and second 0·2 g.mol. oxygen/g.mol. ethyl linoleate respectively. Analyses indicated that at least 90 % of the volatile products was water.

Peroxide formation. Measurement of the peroxide content carried out at the conclusion of or during the course of oxidation runs at 45° C have shown that divergent results are obtained from the two analytical methods—ferrous thiocyanate and Dastur-Lea iodometric—employed. This is illustrated in figure 3, where the fraction (Y) of oxygen atoms in the oxidized linoleate which are active in oxidizing Fe^{++} or HI respectively is plotted against the total extent of oxidation. This represents the first instance encountered where more than half the total reacted oxygen is active towards either ferrous thiocyanate or hydriodic acid (in particular, it may be mentioned that both methods agree in estimating the active oxygen content of oxidation products from the very similar ethyl linolenate at exactly half the total oxygen absorbed). Pending further investigation of the matter, peroxide determinations by the iodometric method have been taken as correct.

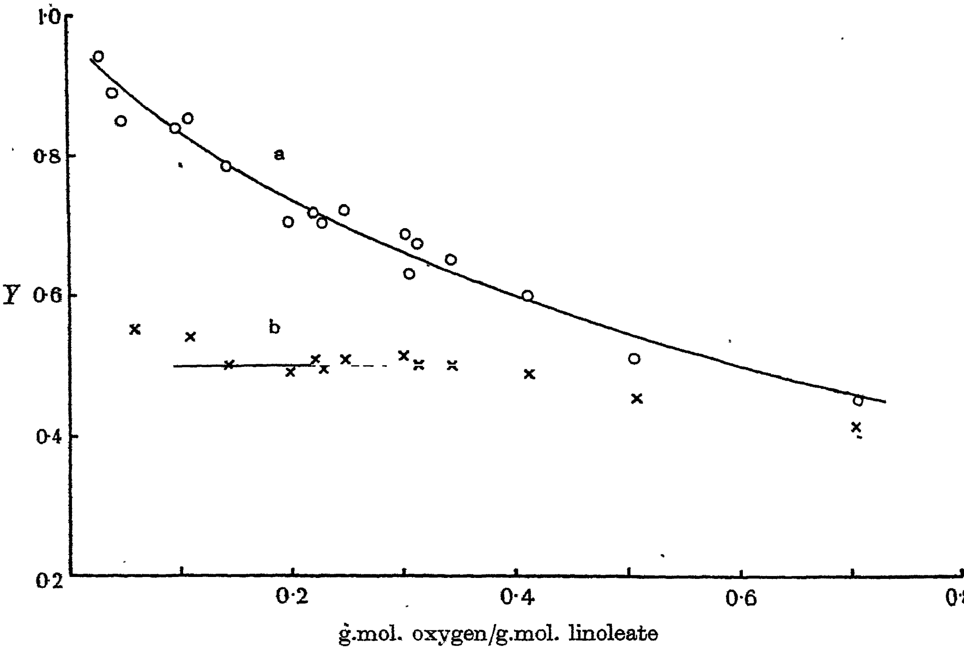


FIGURE 3. Determination of peroxide content of oxidized ethyl linoleate. Y represents the fraction of oxygen atoms in oxidation products active in oxidation of ferrous ion (curve *a*) and KI (curve *b*).

Active hydrogen content. The active hydrogen content of samples of oxidized ethyl linoleate was determined by the method described by Bolland (1941). Duplicate analyses gave the following results: the peroxide contents (based on iodometric estimations and on measured extent of oxidation of samples) of three samples were

0·290, 0·272 and 0·120 g.mol. peroxidic oxygen/g.mol. ethyl linoleate respectively; the corresponding active hydrogen contents were determined as 0·143, 0·141 and 0·062 g.mol. active hydrogen/g.mol. ethyl linoleate. In each case the active hydrogen content was, within experimental error, half the peroxidic oxygen content.

Chromatographic analysis of oxidation products. Samples of oxidation products were dissolved in 40–60° C petrol ether and passed through a short (*ca.* 10 cm.) column of activated alumina (Savory & Moore) in an atmosphere of nitrogen. The chromatogram was eluted with a further quantity of 40–60° petrol ether, which was collected in several separate successive portions. These were evaporated down *in vacuo* at room temperature and the residues weighed. In one experiment a solution of 4·2500 g. oxidized ethyl linoleate, containing 0·184 g.mol. oxygen/g.mol. ethyl linoleate, in 150 c.c. petrol ether was chromatographed in the above fashion. The first 250 c.c. petrol ether passing through the column contained 3·4121 g. residue, the next 100 c.c. 0·0075 g., and a further 100 c.c. 0·0010 g. As all three residues contained only negligible amounts of peroxide, the separation between oxidized and non-oxidized ethyl linoleate was taken to be complete at this stage. The oxidized portion remaining in the chromatogram accordingly amounted to 0·8294 g., containing 1·01 g.mol. oxygen/g.mol. ethyl linoleate. A similar experiment with an oxidized ethyl linoleate sample containing 0·20 g.mol. oxygen/g.mol. ethyl linoleate gave a very similar result in indicating that the average oxygen content of the oxidized molecules was 1·05 g.mol. O_2/g.mol. ethyl linoleate.

The effect of altering the various experimental conditions in turn from their standard values was investigated.

Ethyl linoleate concentration. The linear character of the rate of oxidation *v.* extent of oxidation relation was not altered by dilution of ethyl linoleate with its most closely related saturated ester, ethyl stearate. Further, as figure 4 shows, the slope of the linear curve is proportional to the concentration of ethyl linoleate.

Oxygen pressure. A series of short measurements of the rate of oxidation of a sample of partly oxidized ethyl linoleate (oxygen content *ca.* 0·1 g.mol./oxygen/g.mol. linoleate) was carried out while the oxygen pressure was maintained at a succession of widely varying values in the range 4–760 mm. Any increase in the rate of oxidation arising from an increase in extent of oxidation was corrected for by interspersing the series with measurements at a selected standard pressure (actually 17 mm.). A lower limit was placed on the range of pressures employed by the necessity of ensuring that the rate of diffusion of oxygen into the liquid phase did not become a rate-controlling factor. At the lowest pressure included in the series of experiments, from which figure 5 was constructed, a twofold increase in the rate of shaking the reaction vessel had no effect on the rate of oxidation.

The results of similar sets of measurements at 22 and 65° C are incorporated in table 1.

Temperature. Oxidation runs carried out in the temperature range 35–75° C conformed to the character shown in figure 2. The slopes of the initial portions of the rate of oxidation *v.* extent of oxidation curves are recorded in table 1.

Oxidation of ethyl linoleate in presence of benzoyl peroxide. In a short series of oxidation runs at 45° C, the effect of addition of benzoyl peroxide in various concentrations was determined. In each experiment the oxidation was followed sufficiently far to give an accurate (extrapolated) value of the rate of oxidation at zero extent of oxidation. The initial rates of oxidation conformed accurately to the equation

$$(\text{Rate})_0 = 1{\cdot}20 \times 10^{-3}[C]^{\frac{1}{2}} + 0{\cdot}30 \times 10^{-4}, \tag{2}$$

where $[C]$ represents the gram-molecular fraction of dibenzoyl peroxide (cf. figure 6).

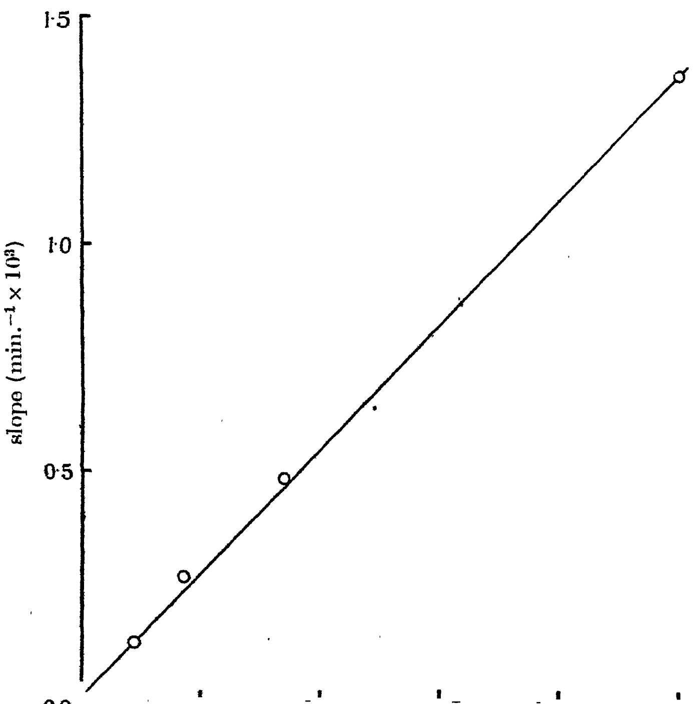


FIGURE 4. Dependence of rate of oxidation on concentration of ethyl linoleate at 45° C and 100 mm. oxygen pressure.

Thermal decomposition of ethyl linoleate hydroperoxide. It will be shown below that the thermal decomposition of the linoleate hydroperoxide plays an important role in the oxidation mechanism suggested, and this reaction has therefore been studied in some detail. A full account will be published elsewhere; it is sufficient for present purposes to note the main conclusions:

(*a*) Decomposition is substantially bimolecular with respect to peroxide concentration, the velocity constant at 45° C, for example, being $1{\cdot}50 \times 10^{-4}$ (g.mol. peroxide/g.mol. linoleate)$^{-1}$ min.$^{-1}$.

(*b*) Not more than one molecule of water is formed for every two peroxide molecules decomposed.

(*c*) The temperature coefficient corresponds to an apparent energy of activation of 26 kcal./g.mol.

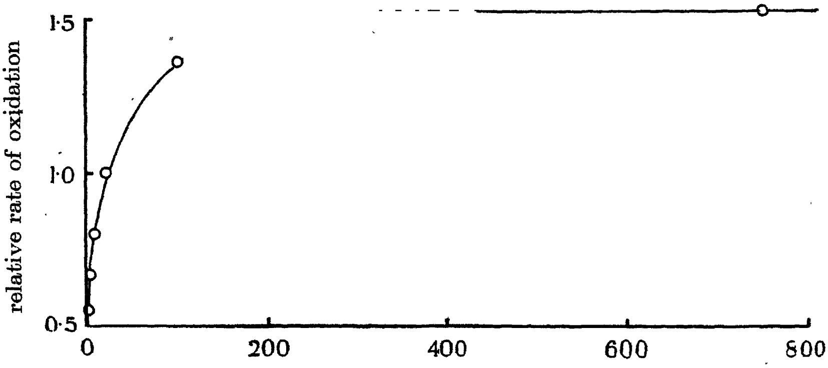


FIGURE 5

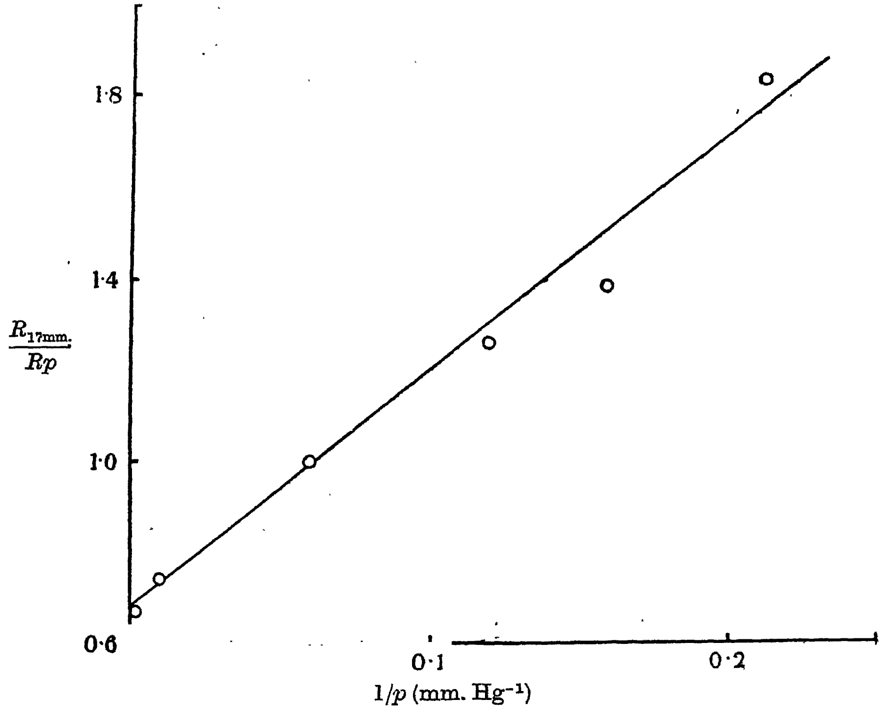


FIGURE 5A

FIGURES 5, 5A. Dependence of rate of oxidation of ethyl linoleate on oxygen pressure at 45° C. $R_{17\,\text{mm.}}$ and R_p represent the rates of oxidation at 17 mm. and p mm. oxygen pressure respectively.

Thermal decomposition of dibenzoyl peroxide. It has been demonstrated that the thermal decomposition of dibenzoyl peroxide in benzene solution is a unimolecular reaction (McClure, Robertson & Cuthbertson 1942). This has been checked under conditions more closely approximating to those under which the oxidation runs were carried out: for this purpose a 6·10 molar % solution of dibenzoyl peroxide in ethyl stearate (reasons of economy precluded the use of ethyl linoleate for this purpose) was heated at 45° C in presence of 10 mm. carefully purified nitrogen (included to prevent evaporation of ethyl stearate). At intervals samples were withdrawn and the peroxide content measured iodometrically (Braun 1942). Over a fourfold diminution in peroxide concentration results consistent with unimolecular decomposition were obtained. The average value of the unimolecular velocity coefficient at 45° C was $3{\cdot}0 \times 10^{-5}$ min.$^{-1}$.

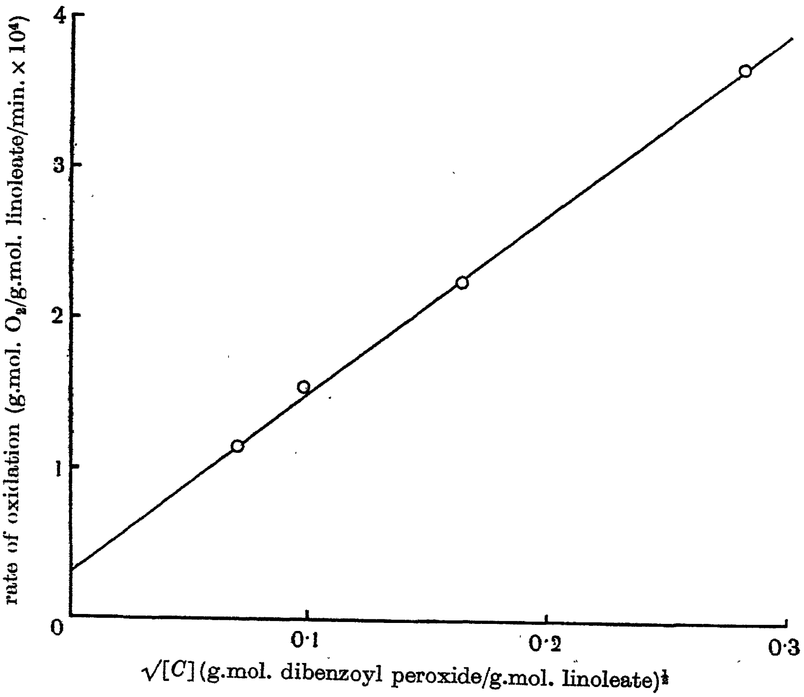
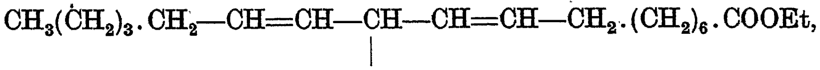
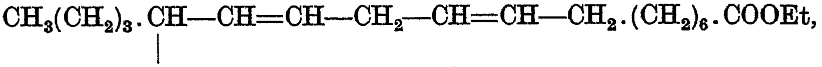


FIGURE 6. Effect of addition of dibenzoyl peroxide on the rate of oxidation of ethyl linoleate at 45° C and 100 mm. oxygen pressure.

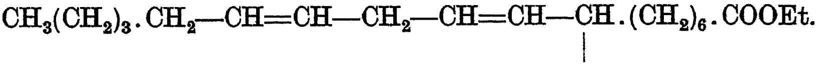
DISCUSSION

In the initial stages of the interaction between molecular oxygen and ethyl linoleate, peroxide groups are formed exclusively: determination of the peroxide content of oxidation products by the Dastur-Lea iodometric method of estimation shows that formation of non-peroxidic products does not become appreciable until some 0·4 g.mol. oxygen/g.mol. ethyl linoleate has reacted. This quantitative aspect of the reaction is emphasized by the almost complete absence of such volatile products as water and carbon dioxide; in sharp contrast, during the thermal oxidation of certain other olefinic materials, such as rubber, which give only fractional yields of peroxide, water and carbon dioxide are formed from the outset.* The active hydrogen estimations quoted on p. 223 show that the peroxide groups formed in the thermal oxidation of ethyl linoleate, as in the case of other olefinic substances (Criegee, Pilz & Flygare 1939; Farmer & Sundralingam 1942, 1943; Farmer & Sutton 1942, 1943; and Sutton 1944) are of the hydroperoxidic type. While no formal proof is offered here that these hydroperoxide groups are disposed in the α-position to a double bond, the observation (Bolland & Koch 1945) that the introduction of a hydroperoxide group may result in the formation of a conjugated diene chromophore in the oxidized molecule represents very strong evidence that the double bonds are not destroyed in the process. The obvious inference is that the alternative reactive points in the molecule, the three α-methylenic groups, constitute the site of reaction. Farmer, Koch & Sutton (1943) have pointed out that the most satisfactory explanation of the diene conjugation which accompanies the oxidation of 1:4 unsaturated esters lies in the formation, at some point in the oxidation mechanism, of free radicals, formally derivable from the parent ester by abstracting a hydrogen atom from an α-methylene group. In the comparatively simple case of ethyl linoleate only three radicals of this type are possible:

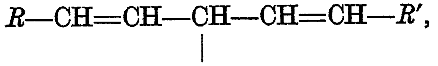

$$CH_3(CH_2)_3.CH_2{-}CH{=}CH{-}\underset{|}{CH}{-}CH{=}CH{-}CH_2.(CH_2)_6.COOEt, \qquad (II)$$


$$CH_3(CH_2)_3.\underset{|}{CH}{-}CH{=}CH{-}CH_2{-}CH{=}CH{-}CH_2.(CH_2)_6.COOEt, \qquad (III)$$

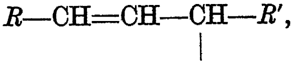
and


$$CH_3(CH_2)_3.CH_2{-}CH{=}CH{-}CH_2{-}CH{=}CH{-}\underset{|}{CH}.(CH_2)_6.COOEt. \qquad (IV)$$

The factor which governs the relative probabilities with which these three radicals are formed must of necessity be the resonance energy made available, by the breaking of the (α-methylenic) carbon hydrogen bond, for stabilizing the resulting free radicals. Calculations of the resonance energies concerned have been made by Orr† (Bolland & Orr 1945), who estimates that the resonance energies associated with the radical systems


$$R{-}CH{=}CH{-}\underset{|}{CH}{-}CH{=}CH{-}R', \qquad (V)$$

and


$$R{-}CH{=}CH{-}\underset{|}{CH}{-}R', \qquad (VI)$$

are 30·5 and 18·7 kcal./g.mol. respectively. On this basis formation of the radical (II) is very much more probable than formation of (III) or (IV). Though the position into which substituents (e.g. —OOH) are introduced into the radical by succeeding reaction processes may well be affected by other factors, it is probably of significance that the formation of products of the types (VII) and (VIII) will be favoured at the expense of (IX) by the resonance energy (*ca.* 7 kcal./g.mol.) associated with their conjugated structure:

---

* Thus if rubber is oxidized in the temperature range 25–100° C, about 0·05 g.mol. H_2O and 0·02 g.mol. CO_2 are formed per g.mol. of oxygen reacted.

† I wish to thank Dr W. J. C. Orr for making available these calculations and for the benefit of discussion on the importance of resonance energies in the present problem.

$$CH_3(CH_2)_4.CH{=}CH{-}CH{=}CH{-}CH(OOH){-}(CH_2)_7.COOEt, \qquad (VII)$$

$$CH_3(CH_2)_4.CH(OOH){-}CH{=}CH{-}CH{=}CH{-}(CH_2)_7.COOEt, \qquad (VIII)$$

$$CH_3(CH_2)_4.CH{=}CH{-}CH(OOH){-}CH{=}CH{-}(CH_2)_7.COOEt. \qquad (IX)$$

It is important to note that formation of radicals is less likely from mono-hydroperoxide molecules than from the unoxidized ester: in the case of the isomers (VII) and (VIII) the resonance energy made available by radical formation is some 23 kcal./g.mol. (i.e. the resonance energy of radical (V) less that of the conjugated diene structure), while in the case of the non-conjugated hydroperoxide (IX) the resonance energy concerned will be that of the allyl radical (VI), viz. 19 kcal./g.mol. It is thus to be anticipated that the oxidative attack on ethyl linoleate mono-hydroperoxide is relatively difficult, so that in the early stages of oxidation the formation of even the most probable di-hydroperoxides, e.g.

$$CH_3(CH_2)_3CH(OOH).CH{=}CH{-}CH{=}CH{-}CH(OOH).(CH_2)_7COOEt \qquad (X)$$

is small. The distribution of the hydroperoxide groups as determined from chromatographic analysis of oxidation products containing up to 0·2 g.mol. O_2/g.mol. linoleate (cf. p. 223) provide confirmation of this view.

On chemical and spectroscopic evidence alone, it thus appears that in the early stage of the thermal oxidation of ethyl linoleate α-methylenic hydroperoxides are formed quantitatively and in restricted variety, and that the intermediate formation of free radicals of type (II) form an essential part of the oxidation mechanism. The kinetic measurements, as will appear in the sequel, are entirely compatible with this general picture, and in addition fill in much detail regarding the oxidation mechanism, in a manner which it is beyond the possibility of purely chemical investigation to provide.

Kinetic analysis

Under all experimental conditions examined, the course of oxidation (cf. figure 2) shows the well-known characteristics of the autocatalytic reaction, but it is to be noted that a finite rate of oxidation is observed even at zero extent of oxidation. This points to the composite nature of the oxidation, and, indeed, examination has shown that it is permissible to attribute the overall rate of oxidation to two distinct types of reaction, which may with advantage be considered independently. The results of the oxidation rate measurements detailed in the experimental section may be summarized by the following expression:

$$\text{rate of oxidation} = \frac{d[R\text{OOH}]}{dt} = K.[R\text{OOH}][R\text{H}].f([\text{O}_2]) + f'([R\text{H}],[\text{O}_2]), \qquad (3)$$

where [RH], [ROOH] and [O_2] represent the concentration of ethyl linoleate, ethyl linoleate hydroperoxide and oxygen in the liquid phase. The first term is identified with oxidation, which being catalysed by the sole product oxidation (i.e. hydroperoxide) is responsible for the autocatalytic aspects of the oxidation and represents the predominant mode of reaction in all but the very initial stages of oxidation. The form of the second, peroxide-independent, term $f'([RH], [O_2])$ has not as yet been completely established. This type of oxidation obviously is of prime importance in the earliest stages of the oxidation of peroxide-free ethyl linoleate in so far as it provides the initial concentration of peroxide, from which the peroxide-catalysed oxidation develops in autocatalytic fashion. In the discussion following, however, attention is concentrated entirely on the peroxide-catalysed oxidation.

Any reaction mechanism which may be advanced for the peroxide-catalysed oxidation of ethyl linoleate must fulfil three essential requirements: (*a*) it must account for the kinetic characteristics of the oxidation, (*b*) it must provide a satisfactory explanation of the catalytic action of peroxide, and (*c*) it must be consistent with information derived from chemical evidence.

In considering many alternatives, the reaction scheme (R 1–R 6) alone has been found to satisfy the first two of these three criteria. This mechanism requires that the thermal oxidation of ethyl linoleate should, in common with many different types of oxidation, be a chain reaction,* and that chains should be initiated by the formation of free radicals from the thermal decomposition of linoleate hydroperoxide, and terminated by the mutual destruction of the particular radicals which act as chain carriers. In the somewhat generalized formulation below, these chain carriers are denoted by X and Y:

$$\begin{array}{llll}
\text{Initiation:} & 2R\text{OOH} \rightarrow X \text{ or } Y & k_1,\ E_1, & (R1) \\
\text{Propagation:} & X + O_2 \rightarrow Y + & k_2,\ E_2, & (R2) \\
 & Y + R\text{H} \rightarrow X + R\text{OOH} & k_3,\ E_3, & (R3) \\
\text{Termination:} & \left.\begin{array}{l} X + X \rightarrow \\ X + Y \rightarrow \\ Y + Y \rightarrow \end{array}\right\} \text{end-products} & \begin{array}{l} k_4,\ E_4, \\ k_5,\ E_5, \\ k_6,\ E_6, \end{array} & \begin{array}{l} (R4) \\ (R5) \\ (R6) \end{array}
\end{array}$$

where k_1 to k_6 and E_1 to E_6 represent the reaction velocity coefficient and energies of activation of the elementary reactions (R 1) to (R 6) respectively.

Two limiting forms which the rate equation deduced from the above mechanism may assume, will be discussed in the first instance. The stationary concentration of the chain carrier, X, will be determined by the rate at which it is formed by the propagation reaction (R 3) and the rate at which it is destroyed by termination reactions ($R4$) and ($R5$); $[Y]$ is controlled in similar fashion by propagation reaction ($R2$) and termination reactions ($R5$) and ($R6$). If the value of $[\mathrm{O_2}]$ is sufficiently high relative to $[R\mathrm{H}]$ (i.e. at sufficiently high oxygen pressure) the stationary concentration of X and Y will be adjusted so that $[X]$ is very large in comparison with $[Y]$. Under these circumstances chain termination will be effected entirely by ($R6$), and the rate equation becomes

---

* Formal confirmation of this view has been obtained from measurements of the effect of inhibitors on the rate of oxidation of ethyl linoleate (J. L. Bolland & P. ten Have, unpublished work).

$$d[R\mathrm{OOH}]/dt = k_3\sqrt{(k_1/k_6)}\,[R\mathrm{OOH}]\,[R\mathrm{H}]. \tag{4}$$

Similarly, when $[\mathrm{O_2}]$ is sufficiently low the stationary concentration of X assumes a value low relative to that of Y, and termination by ($R4$) predominates. Under these conditions

$$d[R\mathrm{OOH}]/dt = k_2\sqrt{(k_1/k_4)}\,[R\mathrm{OOH}]\,[\mathrm{O_2}]. \tag{5}$$

Equation (4) is in complete accord with our experimental results at high oxygen pressures, since in the pressure range 100–700 mm. the rate of oxidation was found to be proportional to the hydroperoxide content (figure 2) and ethyl linoleate concentration (figure 4) and almost independent of oxygen pressure (figure 5). The increasing influence of the oxygen pressure on the rate of oxidation as the former is reduced from 100 to 4 mm. is apparent from figure 5. The intervention of experimental difficulties prevent reliable rate measurements being made below about 4 mm. oxygen pressure, but if it is legitimate to extrapolate the straight line shown in figure 5A it follows that equation (5) is obeyed at still lower oxygen pressures.

In order to derive a kinetic equation applicable at all oxygen pressures, reaction (5) has to be considered as well as (4) and (6). No simple general solution is then possible, but the result is obtained in a very simple form by making the plausible assumption that $k_5^2 = k_4 k_6$:

$$\frac{d[R\mathrm{OOH}]}{dt} = k_3\sqrt{\frac{k_1}{k_6}}\,[R\mathrm{OOH}]\,[R\mathrm{H}]\,\frac{k_2\sqrt{k_6}\,[\mathrm{O_2}]}{k_2\sqrt{k_6}\,[\mathrm{O_2}] + k_3\sqrt{k_4}\,[R\mathrm{H}] + \sqrt{(k_1k_4k_6)}\,[R\mathrm{OOH}]}. \tag{6}$$

When the oxidation chain length is large—as is shown below to be the case in the present system—the term $\sqrt{(k_1k_4k_6)}\,[R\mathrm{OOH}]$ is easily shown to be of quite negligible importance, so that the dependence of oxidation rate on oxygen pressure is of the form found experimentally (cf. figure 5A where 1/rate is shown to be a linear function of $1/p$).

The oxidation chain length, ν, under any particular set of conditions in which the suggested mechanism holds good, is given by

$$\nu = \frac{R}{\alpha R_D} = \frac{1}{\alpha}\,\frac{k_3}{\sqrt{(k_1k_6)}}\,\frac{[R\mathrm{H}]}{[R\mathrm{OOH}]}\,\frac{k_2\sqrt{k_6}\,[\mathrm{O_2}]}{k_2\sqrt{k_6}\,[\mathrm{O_2}] + k_3\sqrt{k_4}\,[R\mathrm{H}]}, \tag{7}$$

where R represents the rate of oxidation, R_D the rate of thermal decomposition of ethyl linoleate hydroperoxide, and α is the average number of fresh chain carriers resulting from the decomposition of one hydroperoxide molecule.

Under 'standard' conditions of temperature (45° C), oxygen pressure (100 mm.) and initial linoleate concentration (100 %), equation (7) reduces to

$$\nu = \frac{9{\cdot}6}{\alpha[R\mathrm{OOH}]}, \tag{8}$$

when $[R\mathrm{OOH}]$ is expressed in g.mol. peroxidic oxygen/g.mol. ethyl linoleate. For example, at extent of oxidation 0·1 g.mol./oxygen/g.mol. ethyl linoleate the chain length is 96, if α is assumed to be unity. It is of interest that the amount of water formed during oxidation is consistent with a chain length of this order. On the basis that one molecule of water is formed for every two peroxide molecules thermally decomposed (cf. p. 224), the average chain length during the introduction of the first and second 0·2 g.mol./oxygen/g.mol. ethyl linoleate is 40 and 55 respectively, compared with values of approximately 96 and 32 calculated from (8).

Oxidation of ethyl linoleate catalysed by dibenzoyl peroxide

The manner in which the thermal decomposition of dibenzoyl peroxide may initiate chain reactions of the radical type (Smith 1945), and the nature of the reaction products with cyclohexene (Farmer & Michael 1942) indicate that the thermal decomposition of dibenzoyl peroxide results in the production of free radicals. While these are almost certainly phenyl C_6H_5- and benzoate C_6H_5COO- radicals, details of the mechanism whereby they are formed are by no means established (McClure *et al.* 1942). If these radicals are capable of initiating oxidation chains in ethyl linoleate, a reaction mechanism differing from that suggested above (R1–R6), only in so far as the initiation step is concerned, is to be anticipated. On this basis, the resulting rate of oxidation will be given by

$$\frac{d[R\mathrm{OOH}]}{dt} = k_3 \sqrt{\frac{k_7}{k_6}}\sqrt{[C]}\,[R\mathrm{H}]\frac{k_2\sqrt{k_6}\,[\mathrm{O}_2]}{k_2\sqrt{k_6}\,[\mathrm{O}_2]+k_3\sqrt{k_4}\,[R\mathrm{H}]}, \tag{9}$$

where $[C]$ and k_7 represent the concentration and the (unimolecular) reaction velocity coefficient for the thermal decomposition of dibenzoyl peroxide.

The underlying reason for the appearance of the square root of the rate of chain initiation in the theoretical rate equations ($\sqrt{k_1}\,[R\mathrm{OOH}]$ and $\sqrt{k_7}\,\sqrt{[C]}$ in equations (6) and (9) respectively) lies in the assumption that chain termination occurs by mutual deactivation of two chain carriers. Accordingly, the proof by direct experiment (cf. equation (2)) that the rate of oxidation catalysed by dibenzoyl peroxide is in fact proportional to $\sqrt{[C]}$ provides valuable confirmation of the correctness of the suggested chain termination mechanism (R4–R6).

When comparing the chain lengths of oxidation, initiated by dibenzoyl peroxide and linoleate hydroperoxide, conditions must be chosen under which the rates of peroxide decomposition are equal. Thus at 45° C the rate of decomposition ($1{\cdot}5 \times 10^{-6}$ g.mol. peroxide/g.mol. ethyl linoleate/min.) is the same in samples of ethyl linoleate containing 0·05 g.mol. dibenzoyl peroxide/g.mol. ethyl linoleate and 0·10 g.mol. linoleate hydroperoxide/g.mol. ethyl linoleate respectively. Since the rates of oxidation (at 100 mm. oxygen pressure) of these samples are $3{\cdot}2 \times 10^{-4}$ and

$1{\cdot}4 \times 10^{-4}$ g.mol. O_2/g.mol. linoleate/min., the respective chain lengths are 213 and 96. In view of uncertainty as to the number and chain-starting efficiency of the radicals formed by the decomposition of the two different peroxides, it is felt that the correspondence between the chain lengths is satisfactory and provides convincing evidence that the thermal oxidation of ethyl linoleate is in fact initiated by hydroperoxide decomposition.

Identity of chain carriers

The foregoing discussion of the oxidation of ethyl linoleate, catalysed by ethyl linoleate hydroperoxide and dibenzyl peroxide, shows that the mechanism (R1–R6) does satisfy the first two of the three requirements given on p. 229; in particular, the nature of the chain-initiation and chain-termination steps have, it is felt, been fully established. It is important to note that no assumptions regarding the actual form that the chain carriers X and Y take have been involved.

In deciding the identity of the chain carriers the remaining requirement—consistency between kinetic and chemical conclusions—is of importance: due weight must be given to the spectroscopic evidence which points conclusively to the formation during the oxidation chain of linoleate radicals (of types (V) and (VI)) which is denoted, in conformity with the other abbreviated formulae, by R—. In order to reconcile this view and the requirement that one propagation reaction must, from kinetic analysis, involve an oxygen molecule and the other an ethyl linoleate molecule, X is identified as R— and Y as RO_2—, i.e. molecules with the radical end-group O—O— substituted in one of the three alternative α-methylene groups.

The precise sequence of reactions by which R— or RO_2— radicals are in the first instance formed from the decomposition of the hydroperoxide has not been determined, but doubtless all the possible free radicals which might be primary products of decomposition could generate R— radicals by the removal of hydrogen atoms from ethyl linoleate molecules.

It is not of course possible to determine the correct chemical formulation of the three termination reactions from kinetic evidence; in the case of (R4) and (R5) the possible alternatives to the formation of R—R and R—O—O—R bridges between linoleate molecules are not obvious. On the other hand, in the third termination reaction (R6) cross-linking must almost certainly be accompanied by the elimination of oxygen or the introduction of oxygenated groups in the neighbourhood of the α-methylene groups concerned. In the early stages of oxidation such cross-linking will be of minor significance; when, however, oxidation reaches more advanced stages, and the chain length presumably becomes short (cf. equation (8)), this type of reaction will become increasingly important and may make a definite contribution towards the 'drying' phenomena which materials like ethyl linoleate exhibit.

An alternative which is formally satisfactory from the kinetic standpoint is that the chain carriers are H— and HO_2—. As this view, however, provides no possible explanation of the conjugated structure of the oxidation products, it is not considered tenable.

The activation energies of the elementary reactions

The apparent energy of activation calculated from the temperature coefficient of the oxidation rate is a composite quantity, and according to (4) equal to $E_3+\frac{1}{2}(E_1-E_6)$ under conditions of very high oxygen pressure. Extrapolation of directly determined rates of oxidation (normally at 100 mm.) to infinite pressure may be carried out by means of the equation

$$R_\infty/R_p = 1+\frac{\gamma}{p}[RH], \tag{10}$$

which, it was noted earlier, is in agreement both with experiment and the approximate theoretical equation (6). Table 1 includes measurements of the constant γ at three temperatures, determinations of the slope of the linear rate of oxidation *v.* extent of oxidation curves in the temperature range 35–75° C, and the corresponding values at infinite oxygen pressure calculated from (10). The energy of activation derived from the last set of figures is 17·2 kcal./g.mol. (cf. figure 7).

TABLE 1

temp. °C	γ mm. Hg	R_∞/R_{100}	slope × 10^3 min.$^{-1}$ 100 mm.	slope × 10^3 min.$^{-1}$ $p=\infty$
22	2·6	1·03	—	—
35	—	1·06	0·58	0·61
45	7·7	1·08	1·36	1·47
54	—	1·12	2·72	3·04
65	18	1·18	6·9	8·2
75	—	1·26	10·5	13·2

From equations (4) and (5) the ratio of rates of oxidation at infinite pressure and a very low oxygen pressure (p) may be written

$$\frac{R_\infty}{R_p} = \frac{k_3}{k_2}\sqrt{\frac{k_4}{k_6}}\frac{[RH]}{\beta p}, \tag{11}$$

where $\beta p = [O_2]$.

This relation is of the same form as the limiting expression derived from (10) when $p \ll \gamma$. If, therefore, the assumption that the linear $1/R$ *v.* $1/p$ relation may be extrapolated to lower pressures than those actually investigated is made, then

$$\frac{k_3}{k_2}\sqrt{\frac{k_4}{k_6}}\frac{1}{\beta} = \gamma, \tag{12}$$

and

$$E_\gamma = E_3-E_2+\tfrac{1}{2}(E_4-E_6)-\Delta H_s, \tag{13}$$

where E_γ is the apparent activation energy calculated from the temperature coefficient γ, and ΔH_s is the heat of solution of oxygen in ethyl linoleate. The value of E_γ obtained from figure 7 is 8·7 kcal./g.mol., and ΔH_s is taken as zero: thus

Horiuti (1933) found negligible values for the heat of solution of oxygen in a series of six organic liquids. Inserting these values in (13)

$$E_3 - E_2 + \tfrac{1}{2}(E_4 - E_6) \approx 8{\cdot}7 \text{ kcal./g.mol.} \tag{14}$$

We also have $E_3 + \tfrac{1}{2}(E_1 - E_6) = 17{\cdot}2$ kcal./g.mol., or inserting the value of E_1 (26 kcal./g.mol.) deduced from the kinetics of the thermal decomposition of ethyl linoleate hydroperoxide,

$$E_3 - \tfrac{1}{2}E_6 = 4{\cdot}2 \text{ kcal./g.mol.} \tag{15}$$

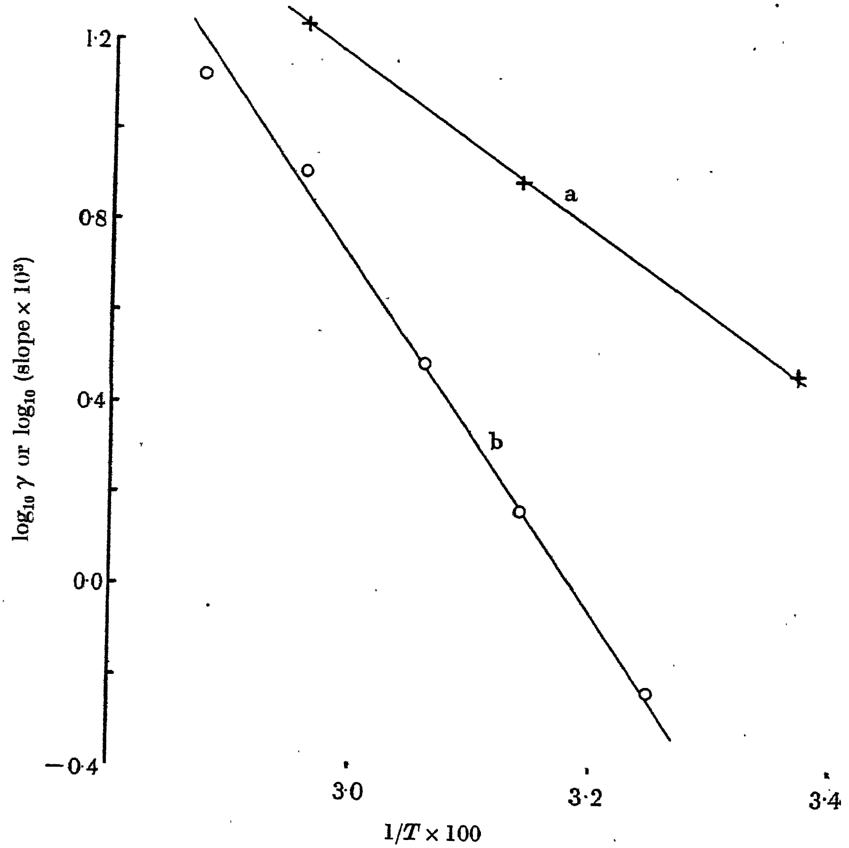


FIGURE 7. Dependence on temperature of γ (curve a) and the slope of the rate of oxidation v. extent of oxidation relation (curve b).

It is obviously not possible to evaluate the four activation energies involved from (14) and (15). E_4, since it corresponds to recombination of hydrocarbon radicals, may, however, be taken to be about 6 kcal./g.mol. On this assumption, E_2 is estimated at −1·5 kcal./g.mol. E_3 and E_6 are interrelated in accordance with (15). Without making any assumptions as to the magnitude of E_6 (which may well be different from E_4) we may therefore estimate E_3 at not less than 4·2 kcal.

In view of uncertainty as to some of the energies introduced in the above argument (especially E_γ and ΔH_s), it is not perhaps surprising that the value of E_2 so derived is negative. It seems, however, permissible to conclude that the energy of activation required for

$$RO_2— + RH \rightarrow ROOH + R— \qquad (R3')$$

is significantly in excess of that for

$$R— + O_2 \rightarrow RO_2—. \qquad (R2')$$

The thermochemistry of the elementary reactions

A further test which must be applied to the individual reactions which comprise the suggested reaction mechanism is their soundness from the energetic point of view: thus the propagation and termination reactions, to which only small energies of activation have been assigned above, would require to be exothermic or only very slightly endothermic. Rough estimates of these heats of reaction can readily be made from the bond energies appropriate to the various reactants and products. The bond energies employed here are carbon-hydrogen 99 kcal. (Baughan 1941; Kynch & Penney 1942), carbon-oxygen 87 kcal. (Baughan, Evans & Polanyi 1941) and oxygen-hydrogen 110 kcal. The most reliable estimate at present available of the oxygen-oxygen bond strength in hydroperoxides ($D_{O—O}$) is that calculated from the data of Stathis & Egerton (1940) on the heats of combustion of ethyl and propyl hydroperoxides, viz. 56 kcal.; preliminary measurements of the heat of oxidation of ethyl linoleate and ethyl linolenate provide confirmation of this value.

The termination reactions will clearly be exothermic. The condition that the propagation step ($R2'$) should be exothermic is $D'_{O—O} > 54$ kcal., where $D'_{O—O}$ represents the oxygen-oxygen bond energy in the peroxide radical, RO_2. This requirement will obviously be fulfilled provided $D'_{O—O}$ is not significantly less than $D_{O—O}$. $D'_{O—O}$ does in fact exceed $D_{O—O}$ if we accept Pauling's (1940) treatment of the electronic structure of the peroxide ion, on the analogy of which the RO_2 radical may be described as $R—O\overset{\cdots}{—}O$, the three-electron bond serving to make $D'_{O—O} > D_{O—O}$; Pauling's quantitative interpretation of the three-electron bond would suggest that $D'_{O—O} \sim \frac{3}{2} D_{O—O}$.

Since it is unlikely that $D'_{O—O}$ will exceed $D_{O—O}$ by more than 41 kcal., the energetic requirement for the second propagation reaction ($R3'$) is considered fulfilled.

The final conclusion is that the set of elementary reactions which permit of the complete interpretation of the kinetic data is also satisfactory from the thermochemical point of view.

I wish to express my thanks to Dr G. Gee for the benefit of many helpful and stimulating discussions during the course of the above work, which forms part of the programme of fundamental research undertaken by the Board of the British Rubber Producers' Research Association.

REFERENCES

Baughan, E. C. 1941 *Nature*, **147**, 542.
Baughan, E. C., Evans, M. G. & Polanyi, M. 1941 *Trans. Faraday Soc.* **37**, 377.
Bolland, J. L. 1941 *I.R.I. Trans.* **16**, 267.
Bolland, J. L. & Koch, H. P. 1945 *J. Chem. Soc.* p. 445.
Bolland, J. L. & Orr, W. J. C. 1945 *I.R.I. Trans.* **21**, 133.
Bolland, J. L., Sundralingam, A., Sutton, D. A. & Tristram, G. R. 1941 *I.R.I. Trans.* **17**, 29.
Braun, G. 1942 *Organic Syntheses*, **13**, 88.
Criegee, R., Pilz, H. & Flygare, J. 1939 *Ber. dtsch. chem. Ges.* **72**, 1799.
Dastur, N. N. & Lea, C. H. 1941 *Analyst*, **66**, 90.
Farmer, E. H. 1942 *Trans. Faraday Soc.* **38**, 340.
Farmer, E. H., Koch, H. P. & Sutton, D. A. 1943 *J. Chem. Soc.* p. 541.
Farmer, E. H. & Michael, S. E. 1942 *J. Chem. Soc.* p. 513.
Farmer, E. H. & Sundralingam, A. 1942 *J. Chem. Soc.* p. 121.
Farmer, E. H. & Sundralingam, A. 1943 *J. Chem. Soc.* p. 125.
Farmer, E. H. & Sutton, D. A. 1942 *J. Chem. Soc.* p. 139.
Farmer, E. H. & Sutton, D. A. 1943 *J. Chem. Soc.* pp. 119, 122.
Hershberg, E. B. & Huntress, E. H. 1933 *Industr. Engng Chem.* (Anal. ed.), **5**, 144.
Horiuti, J. 1933 *Z. Elektrochem.* **39**, 20.
Kohman, G. T. 1929 *J. Phys. Chem.* **33**, 226.
Kynch, G. T. & Penney, W. G. 1942 *Proc. Roy. Soc.* A, **179**, 214.
McClure, J. H., Robertson, R. E. & Cuthbertson, A. C. 1942 *Canad. J. Res.* B, **20**, 103.
Pauling, L. 1940 *The Nature of the Chemical Bond.* Cornell University Press.
Smith, R. R. 1945 *J. Soc. Chem. Ind., Lond.* (in the Press).
Stathis, E. C. & Egerton, A. C. 1940 *Trans. Faraday Soc.* **35**, 606.
Sutton, D. A. 1944 *J. Chem. Soc.* p. 242.

Recombination and the long duration of the Balmer spectrum

By H. ZANSTRA, PH.D., *Howard College, Durban*

(*Communicated by Lord Rayleigh, F.R.S.—Received* 20 *February* 1945)

Lord Rayleigh has recently found that the time of relaxation of the Balmer series of hydrogen excited by an electrodeless discharge is at least of the order 10^{-5} sec. after it has been blown out of the electric field by its own expansion. Following Page, it is assumed that the luminosity is due to recombination of electrons with protons, and then Kramers's theory yields half-value periods certainly larger than 10^{-5} sec. under the special experimental conditions. Difficulties in the application of this theory raised by Cillié and Page are briefly discussed.

The purpose of the present paper is to discuss an experiment by Lord Rayleigh (1944). He finds that if hydrogen made luminous by a powerful discharge is blown out of the electric field by its own expansion, the time of relaxation of the intensity of the Balmer lines is under some conditions of the order 10^{-5} sec. at least. This is about a thousand times more than the life usually adopted for the excited states from which the Balmer series originates, which is of the order 10^{-8} sec. This fact is presented by him for comment.

In the following the decay is worked out under the supposition that the spectrum is produced by recombination. In the case of galactic nebulae, this mechanism of luminosity has been used for explaining the production of the Balmer series and the continuous spectrum at its head by capture of electrons by protons on the various energy levels and subsequent coming down to lower levels. In the laboratory, Page (1938) has investigated the recombination spectrum of hydrogen in the electrodeless discharge. His work does not seem to have been published in full, but was briefly discussed by him a few years ago in a letter to *Nature*. Unfortunately, the title 'Kramers's absorption law in astrophysical problems' made no reference to his experiments, so that it may have been overlooked easily by those interested in the experimental side. Page measured the spectrum in the electrically undisturbed region, and states that recombination of protons and electrons must account for almost the entire emission.

It would seem that the same applies to Rayleigh's experiment. Where the hydrogen is blown out of the electric field there is no direct excitation, and one may expect the excited states to be chiefly produced by capture of electrons on the energy levels of the protons. Besides this there may be the collision excitation due to the free electrons, but this, as will be shown, is only of short duration. Our considerations will be somewhat simplified, because a possible influence of molecular hydrogen is disregarded.

At any instant t, let N atoms per c.c. be ionized, so that there are N free electrons and N protons. The number of recombinations/c.c./sec. on any level n is then given by N^2C_n, where C_n is a function depending on n of the absolute electron temperature T, assuming the electrons to have a Maxwell distribution. The electron temperature will be considered constant, then C_n is a constant depending on the level. The number of recombinations/c.c./sec. on all levels should be equal to the rate of decrease of N, so that

$$\frac{dN}{dt} = -N^2C, \tag{1}$$

where

$$C = \sum_{n=1}^{\infty} C_n. \tag{1a}$$

C is the recombination coefficient.

Integrating this equation, then

$$\frac{1}{N} - \frac{1}{N_0} = Ct, \tag{2}$$

where N_0 represents N for $t = 0$.

The number of recombinations/c.c./sec. on any level n is then given by

$$L = -N^2C', \tag{3}$$

where

$$C' = \sum_{n=3}^{\infty} C_n. \tag{3a}$$

This also represents the number of quanta emitted/c.c./sec. in the Balmer series of hydrogen or luminosity L, as has been shown elsewhere (Zanstra 1927).*

The luminosity has gone down by a factor 2 when $N^2 = \frac{1}{2}N_0^2$ or $N = N_0/\sqrt{2}$. This occurs in the time τ,

$$\tau = \frac{\sqrt{(2)}-1}{CN_0} = \frac{0{\cdot}414}{CN_0}, \tag{4}$$

as follows from (2).

The decay is not exponential. From (3) it follows that the luminosity has gone down to $\frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{8}$ of its initial value, if N has gone down to $1/\sqrt{2}$, $1/\sqrt{4}$, $1/\sqrt{8}$ of N_0, so that the subsequent half-value periods are τ, $\tau\sqrt{2}$, $\tau\sqrt{4}$, $\tau\sqrt{8}$, etc. Assuming constant electron temperature, the dying down should therefore become slower as time goes on. This effect, however, might be opposed if the electron temperature decreases with the time making for faster recombination.

Besides, by recombination, the free electrons might also produce luminosity by collisions with neutral atoms. Since initially their energy cannot be greatly in excess of the ionization energy, the total luminosity of this collision excitation integrated over the time cannot greatly exceed that of recombination. Initially many electrons of the Maxwell distribution still have energies larger than the excitation energy of some hydrogen levels, and then by theory and experiment the effective area of collision is of the order of that of a Bohr orbit 10^{-16} cm.^{2}, whereas the effective area of capture according to Kramers's theory is of the order 10^{-21} cm.^{2} for electrons of such energy. Initially this collision excitation is therefore predominant, but because of this the main kinetic energy is spent in a very short time and leaves only the much slower process of recombination which has just been considered.

In Rayleigh's experiment the pressure was about 0·2 mm., leading to a concentration of $2{\cdot}38 \times 10^{-8}$ g./c.c. or a number of atoms per c.c. of $1{\cdot}44 \times 10^{16}$.

From the heat developed, the energy produced in the gas by one discharge was found to be 10^7 ergs or $0{\cdot}629 \times 10^{19}$ eV. Dividing this by the volume 79 c.c. of the tube and the ionization energy of the hydrogen atom 13·53 eV., one finds that the maximum number of ion pairs which can be produced per c.c. is $0{\cdot}589 \times 10^{16}$, so that at most about four-tenths of the hydrogen can be ionized. One may therefore assume

$$N_0 < 0{\cdot}589 \times 10^{16}, \tag{5}$$

but presumably of this order.

Up to here no special assumptions have been made as regards the form of the coefficient C_n as a function of n and T. The recombination, according to Kramers's theory, will now be introduced (Kramers 1923). It was worked out by Cillié (1932) in connexion with nebulae.* The quantities C_n for this theory are given by

---

* In particular, pp. 56 and 57 of the reference. From equation (9) $Ly_c + Ba_c + Ba = N_{ul}$ follows $Ba = N_{ul} - Ly_c - Ba_c$. N_{ul} may be considered as the number of recombinations on all levels, Ly_c as those on the first and Ba_c as those on the second level, so that the total number of quanta Ba emitted in the Balmer series equals the number of recombinations on levels 3 and higher. The equation presupposes that a quantum of the Lyman series can frequently be absorbed on its way out of the gas, which results in the transformation into the first line Ly_α and other quanta.

$$C_n = 3{\cdot}208 \times 10^{-6} M(n, T), \tag{6}$$

where
$$M(n, T) = \frac{1}{T^{\frac{3}{2}}} \frac{1}{n^3} e^{x_0} \int_{x_0}^{\infty} e^{-x} \frac{dx}{x}, \tag{6a}$$

and
$$x_0 = \frac{\chi_n}{kT}. \tag{6b}$$

χ_n represents the ionization energy for the nth level and k the Boltzmann constant. The rather involved constant in Cillié's formula (6a) has been replaced by its numerical value. In his table 1 Cillié has given the values of $M(n, T)$ for levels up to the tenth. From these and (6) the quantities C of (1a) and C' of (3a) have been computed for the same electron temperatures as Cillié's by summation of the quantities $M(n, T)$. The contribution by levels higher than the tenth is not negligible and was allowed for by replacing the summation for those levels by the integral $\int_{n=11}^{\infty} M(n, T)\, dn$. For the various temperatures between 1000 and 50,000° C of the table it amounts to the percentages 20, 11, 8, 6 and 4 of the total for all levels. The details of this computation need not be given. These results are given in table 1. In each case the last figure of the number given may be somewhat in error.

TABLE 1. VALUES OF C, C' and τ IN SEC. (KRAMERS'S THEORY)

T ...	1000° C	5000° C	10,000° C	20,000° C	50,000° C
$10^{14}C$	218	74	46	27	13·5
$10^{14}C'$	122	33	17·7	9·4	3·6
$10^{5}\tau >$	3·23	9·5	15·3	26	52

The upper limit (5) for N_0 in Rayleigh's experiment and C from the second row substituted in (4) yield the lower limits for the half-value period τ of the dying down luminosity given in the last row. From the heat developed Rayleigh computed a sudden rise in temperature for each discharge of about 35,000° C if the gas is not ionized. Ionization and excitation without collisions of the second kind would lower this temperature which refers to kinetic energy of molecules or neutral atoms. The electron temperature T, however, is determined by quite different factors, and it would be hard to make an *a priori* estimate. An argument for a low electron temperature might be obtained from Page's work. From the relative intensity distribution of the continuous spectrum at the head of the Balmer series, Ba_c, Page derived electron temperatures of about 1000° C both for certain nebulae (Page 1936) and in the electrodeless discharge in the laboratory (Page 1938). If this Ba_c spectrum does not spread far into the ultra-violet, the electron temperature thus determined mainly depends on the general relation between probabilities of absorption and emission for the second level and not on any special type of theory. Assuming for Rayleigh's experiment $T = 1000°$ C, Kramers's theory would predict $\tau > 3{\cdot}23 \times 10^{-5}$ sec. according to the last row of the table, which is of the order observed by Rayleigh. One would scarcely allow an electron temperature much smaller than 1000° C, the minimum being room temperature. If one admits higher values than 1000° C, the half-value period is even larger as is seen in the last row. Whether Kramers's theory be quite correct or not, the conclusion seems justified that recombination would lead to a half-value period much larger than the life of excited states which is of the order 10^{-8} sec., and that the large times of relaxation found by Rayleigh are to be explained by the assumption that the luminosity outside the exciting field is due to recombination, in agreement with Page.

---

* The probability of capture for a level n may be found from the absorption coefficient of radiation for this level by means of a relation between probabilities of absorption and emission found by Kramers and Milne independently. Gaunt's wave-mechanical absorption coefficient for a level differs from that of Kramers only by a factor of the order unity. Since Cillié took the Gaunt factor equal to unity, his Gaunt-Milne treatment is equivalent to that of Kramers and close to that required by wave mechanics.

After having stressed the recombination according to Kramers's theory as a guide for understanding the matter under discussion, it is only fair to mention some difficulties. Cillié (1936) has pointed out that the intensity ratio of the Ba_c to the Balmer spectrum in certain planetary nebulae observed by Page (1936) is very much smaller than predicted by theory and may even be of a different order. This would mean that the share of captures coming to the second level is much smaller than the theory requires. Page (1938) has shown that the same difficulty occurs in his laboratory spectrum where the second level only gets 1/2000 of its theoretical share in the captures. Besides this there are lesser deviations, as Page shows.

One might be tempted to jump to the conclusion that the captures on the second level do not obey Kramers's theory, but Pannekoek's discussion of Menzel & Cillié's data on the chromosphere seems to show that this level obeys the theory fairly well (Pannekoek 1938; Menzel & Cillié 1935, 1937). Pannekoek uses the merging of the higher Balmer lines—distorted by Stark effect—into a continuous spectrum. To produce the required Stark effect, an electron and ion concentration

$$N_e = N_i = 6 \times 10^{11}/\text{c.c.}$$

is required. On the other hand, using Kramers's theory, Menzel & Cillié derive an electron temperature of about 10,000° C from the relative Ba_c distribution and a proton and electron concentration $N_i = N_e = 4 \times 10^{11}$ from the absolute Ba_c intensity. Since the Stark effect is well established experimentally, Pannekoek's determination seems well founded, and the fact that nearly the same concentration gives a good fit with Kramers's theory seems to show that this theory accounts quantitatively for the captures on the second level resulting in Ba_c emission.

One may express the hope that the effect found by Lord Rayleigh will provide another powerful approach to this problem of recombination which as yet seems to be so full of contradictions.

REFERENCES

Cillié, G. C. 1932 *Mon. Not. R. Astr. Soc.* **92**, 820.
Cillié, G. C. 1936 *Mon. Not. R. Astr. Soc.* **96**, 771.
Kramers, H. A. 1923 *Phil. Mag.* **46**, 836.
Menzel, D. H. & Cillié, G. C. 1935 *Harvard Circular*, no. 410.
Menzel, D. H. & Cillié, G. C. 1937 *Astrophys. J.* **85**, 88.
Page, T. L. 1936 *Mon. Not. R. Astr. Soc.* **96**, 604 or *Commun. Univ. Obs. Oxford*, no. 7.
Page, T. L. 1938 *Nature*, **141**, 1137.
Pannekoek, A. 1938 *Mon. Not. R. Astr. Soc.* **98**, 694.
Rayleigh, Lord 1944 *Proc. Roy. Soc.* A, **183**, 20.
Zanstra, H. 1927 *Astrophys. J.* **65**, 50.

The emission of light from spark discharges

By J. D. Craggs and J. M. Meek

(*Communicated by C. R. Burch, F.R.S.—Received 5 March 1945*)

[Plates 7–11]

The investigation is concerned with the transitory effects occurring in spark channels, after the discharge gap has been bridged completely by a streamer. Observations have been made for various gases, in particular for hydrogen and argon, at pressures of the order of 1 atm. The main features studied are the light emitted from the spark channel during the period of current flow, and also the after-glow which persists after the current has fallen to a negligible value. The results indicate that the after-glow in hydrogen is probably a thermal effect only, whilst in argon it is due also, indeed largely, to other causes of which the persistence of atoms in metastable states is the most likely.

Two main experimental methods were employed: (*a*) a revolving mirror camera, similar in principle to those used previously by several investigators, and (*b*) a photoelectric electron multiplier tube coupled directly to a cathode ray oscillograph. The latter method, which is new in investigations of this character, enables quantitative light emission results to be obtained, and methods of calibrating the apparatus are described in detail. The overall time constant of the circuit was sufficiently small to enable after-glows of duration as low as 1μsec. to be clearly distinguished.

The measurements show that for currents of about 100 amp., lasting for 2–4μsec., the after-glow in argon at a pressure of about 1 atm. could be detected after some 30μsec., whereas in hydrogen, for similar conditions, the after-glow lasted for only some 3μsec. Other gases showed after-glows of durations varying between the limits set by hydrogen and argon.

In order to correlate the light output with energy dissipation in the spark channel, calorimetric measurements were made from which the mean voltage drop during the passage of current was estimated. The channel radii were measured with photographic plates sensitive to the visible and near ultraviolet light. Observations were also made of the spectra of the light emitted from the channel. The argon spectrum showed a strong continuum, and, for hydrogen, only the Balmer lines, much broadened, were seen.

The density of ionization in the spark channels is deduced approximately in several ways, from the Stark effect, in hydrogen, and from a consideration of the energy balance in the channel, in both hydrogen and argon. The various calculations are in fairly close agreement and give $N_i \sim 10^{17}$ ions per c.c. Channel temperatures, as determined on the basis of Saha's equation, the limited applicability of which is discussed, are shown to be about 10,000–15,000° K.

The mechanism of light emission from the channel is discussed in some detail, and it is shown that either normal excitation or electron-ion recombination could be entirely responsible for the observed effects.

1. Introduction

The properties of the main spark channel, to be distinguished from the electron avalanche and streamer processes which lead to its formation in a spark discharge (Loeb & Meek 1940), have been little studied. The spark channel may be defined as the conducting path formed, during a spark discharge, after the complete bridging of a gap by a streamer. The physical form of the discharge corresponds to a high-pressure glow or incipient arc and, after a suitable time, if the external circuit conditions are favourable, leads to the formation and maintenance of a stable arc. The present experiments were undertaken partly to obtain information on the mechanism of light emission from the spark channel in various gases and partly to study after-glows.

The present work differs from that of previous authors in that the light measurements are made by a combined photoelectric and oscillographic method and refers not only to the light emitted during the discharge, but also to the after-glow persisting in the gas after the cessation of current flow.

2. Experimental methods

The sparks were observed through about 3 mm. of hard glass, which absorbed most of the radiation for $\lambda < 3300$ A, by means of a synchronized rotating mirror or an electron multiplier photocell directly coupled to a synchronized cathode-ray oscillograph. The current was supplied either from a condenser or, when square pulses were required, from artificial transmission lines (Kennelly 1917). The gases used were of industrial purity, i.e. $\not> 2\,\%$ impurity.

With a load resistance of 10,800 ohms, the time constant of the multiplier output circuit was $0{\cdot}3\,\mu$sec. The multiplier was calibrated for varying intensities of light with a series of calibrated gauzes, using spark sources of light to simulate working conditions. Figures 1 and 2 give typical calibrations, the shape of which were not altered by using light of various colours.

The colour response (figure 4) was determined with a series of glasses and Wratten filters, whose spectral ranges taken with a copper arc are shown in figure 3 (plate 7) and a mercury discharge lamp, calibrated against a special photocell of uniform colour response. The variations shown in figure 4, which represents results taken over 6 months, were probably due partly to small charges in multiplier colour response and to the unsuitability of the emission from the mercury lamp which gave only a small light output for some of the filter combinations.

Finally, absolute calibrations with a further set of calibrated filters and standard lamp were taken. At a source/multiplier cathode distance of 22·5 cm., 0·60, 5·76, 1·96 and $1{\cdot}58\,\mu$W/sq.cm./m. at wave-lengths of respectively 3680, 4600, 4900 and 5330 A, gave multiplier output currents of 0·035, 1·32, 0·51 and 0·38 mA.

As an example, the multiplier sensitivity at 4861 A (H_β) was determined, using the results for the 4900 A filter and correcting the multiplier response. 5 cm. oscillo-

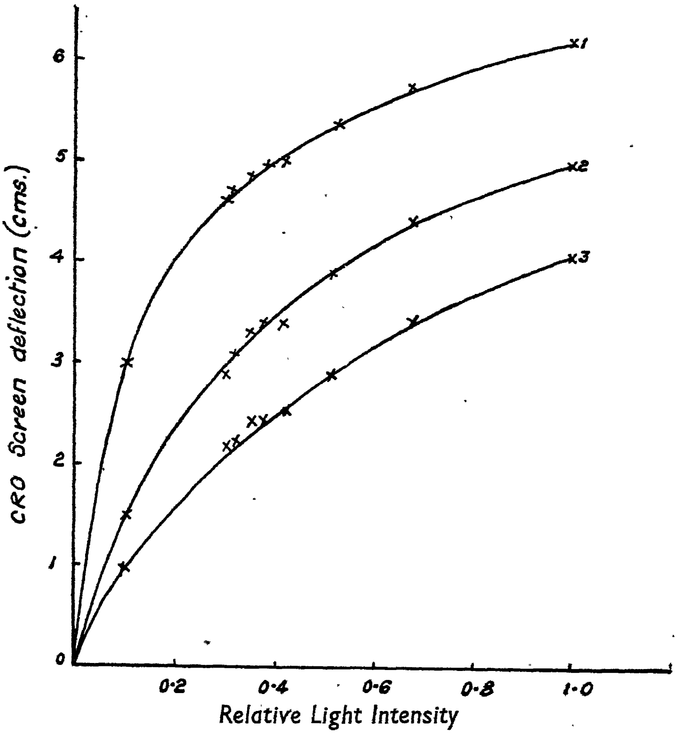


FIGURE 1. Calibration curves for electron multiplier with different values of maximum C.R.O. screen deflexion.

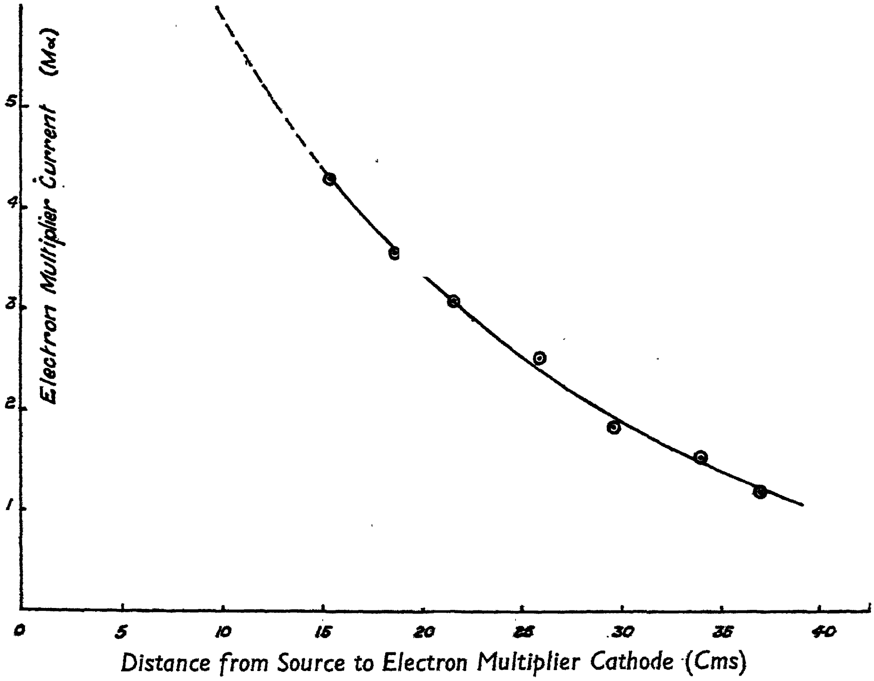


FIGURE 2. Variation of electron multiplier output current with distance from a steady source of illumination to the multiplier cathode.

graph screen deflexion was found to correspond to $1{\cdot}6\times10^{16}$ quanta/sec. at 4861 A. Therefore the total emission from a spark at 15 cm. distance necessary to give that deflexion was $3{\cdot}6\times10^{19}$ quanta/sec.

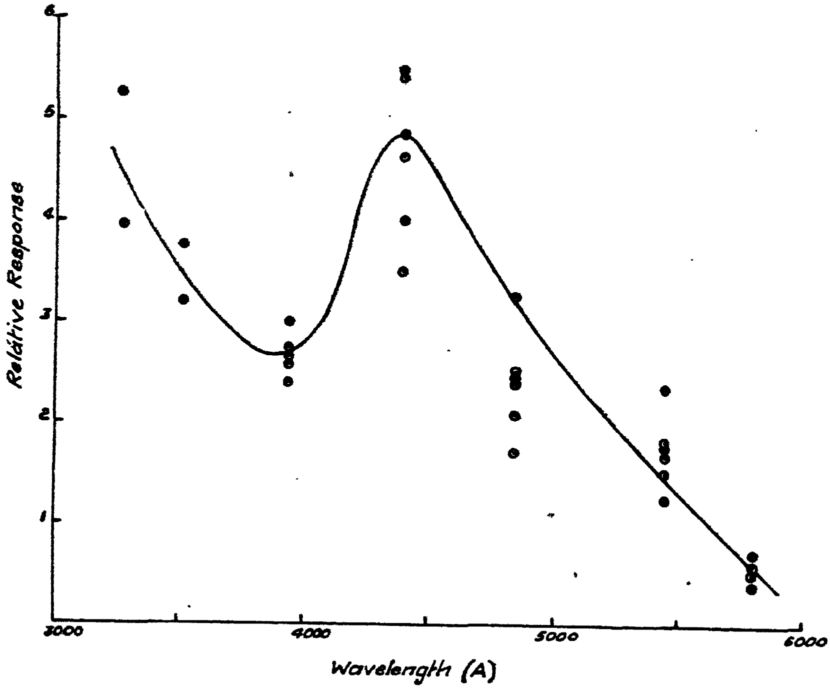


FIGURE 4. Variation of relative response of electron multiplier with wavelength of incident light.

3. EXPERIMENTAL RESULTS

3·1. *Rotating-mirror photographs*

Photographs of spark discharges across a 1 cm. gap between $\frac{1}{16}$ in. diameter tungsten electrodes in hydrogen, argon, nitrogen and helium,. as recorded with the rotating-mirror camera using Ilford Astra III plates, are shown in figure 6 (plate 8). Several 'still' photographs, obtained with the mirror stationary, are also included. In all these records the sparks were produced by the discharge of a $0{\cdot}04\,\mu$F condenser, in series with a 13-ohm resistance. Oscillographic observations showed the resultant current wave-form to be critically damped, and enabled the peak value of the current to be measured. The magnitude of the current fell to less than one-thousandth of the peak value, within 5 μsec. after the initiation of the spark and, as later experiments for argon (figure 12*a*) showed, the light emission varies approximately as the current strength, so that such low currents would have a negligible effect in the maintenance of the observed after-glow in the different gases used if the same mechanisms of light production obtain both during the current pulse (when figure 12*a* applies) and the after-glow. The same mechanisms would not apply if the flow of current affects the light output by virtue of the drift velocity of electrons or ions, because such effects would rapidly disappear after the cessation of current, as simple calculations indicate. However, the channel temperatures and therefore the electronic random velocity are so high that thermal processes, which are operative both during and after the current flow, probably exert a controlling influence. It is also for further reasons considered that the very small residual currents are not responsible for the after-glow. Later experiments (figure 9, plate 9) showed that currents of very marked oscillatory character show relatively little disturbance of after-glow. The great difference in after-glow lengths in argon and hydrogen, where residual currents should be almost equal as they are largely controlled by circuit conditions, indicates that some mechanism other than any caused by such small residual currents, exerts a controlling influence. Other details of the conditions associated with the discharges shown in figure 6 (plate 8) are given in table 1.

TABLE 1

figure	gas	pressure lb./in.^{2} abs.	condenser voltage kV	peak current amp.	photograph
5*a*	H_2	22	12	740	moving
b	H_2	22	12	740	still
c	Ne	22	4·5	280	moving
d	Ne	22	4·5	280	still
e	He	22	6	370	moving
f	He	22	6	370	still
g	N_2	15	12	740	moving
h	N_2	15	12	740	still
i	A	22	7	430	moving
j	A	22	7	430	still
k	A	22	7	430	moving
l	A	22	7	430	moving
m	A	22	8·5	610	moving
n	A	50	12·5	770	moving
o	A	50	12·5	770	moving
p	A	22	8	580	moving
q	A/H_2	19/3	11	680	moving

The hydrogen sparks, three of which are shown in the rotating-mirror photographs of figure 6*a*, give no detectable after-glow, and it is not possible to distinguish any differences between such photographs and those obtained with a stationary mirror.* The spark channel is more sharply defined than that observed in the other gases used and has a diameter of about 0·15 cm., so that the peak current density is of the order of 40,000 amp./cm.^{2}.

In argon the after-glow is appreciable, and persists for times as long as 50 μsec. at the higher currents used. Considerable differences occur between the after-glows for sparks with the same circuit conditions, as shown by comparison of the various records *i* to *p* in figure 6. The after-glow may be either uniform and wedge-shaped on the photographs, with a greater persistence of illumination in the centre of the gap, as in figure 6*m*, or may be composed of a number of more or less separate luminous regions (figure 6*n*).

---

* § 3·2 describes results with higher resolving powers.

The rotating mirror records for nitrogen and helium show no detectable after-glow, for the degree of resolution used. In neon, however, there is a visible after-glow of about 10 μsec., as shown by comparison of the moving and still photographs of figure 6*c* and *d*.

3·2. *Oscillograms of light emission*

The rotating mirror photographs of the discharges are confused to some extent by the difficulty of distinguishing between spatial changes in the channel and the time variations. Such difficulties can be reduced by an increase in the resolving power of the camera, but a more convenient and direct method is afforded by the use of the photoelectric recording system.

A series of oscillograms showing the variation of light with time are shown in figure 5 (plate 7). These refer to discharges of 0·5 cm. length in argon and hydrogen, and in mixtures of these gases, when subjected to a square current pulse of 120 amp. flowing for about 4·5 μsec. An oscillogram of the current pulse is given in figure 5*f*. The total gas pressure was 18 lb./in.^{2} abs., and was constituted in the proportions indicated in table 2.

TABLE 2

	figure 5*a*	figure 5*b*	figure 5*c*	figure 5*d*	figure 5*e*
A	1·0	0·97	0·94	0·75	0
H_2	0	0·03	0·06	0·25	1·0

The light emitted by the hydrogen spark follows fairly closely the shape of the current pulse, and falls rapidly to zero once the current is extinguished. In argon the light reaches a maximum value during the early stages of the spark and then falls, even though the current is maintained at the same value. When the current ceases, an appreciable after-glow lasting many microseconds is recorded in argon. The peak light output for each spark is roughly constant, but the after-glow varies considerably from spark to spark as shown in figure 5*a*, which is a composite oscillogram of the light pulses from a number of sparks. The addition of a small amount of hydrogen to the argon reduces the variability of the after-glow, which tends to stabilize at a higher rather than a lower value.

Analyses of typical light oscillograms for discharges in argon and hydrogen are given in figure 7*a* and *b* respectively. The rate of increase of light in hydrogen is lower than that in argon, and is too low to be accounted for by consideration of the frequency response of the light-recording system. It may partly be explained in terms of the difference of diffusion rates for hydrogen and argon, in that the discharge channel, which is constricted in an abnormal manner, with high current density and light output, requires a longer time to establish maximum light output in hydrogen than in argon. A further consideration affecting the response of the recording system is the change in spectral distribution of the light emitted during the initial and final stages of the spark. Finally, the high specific heat of hydrogen (about 6 times that of argon) tends to impede the formation of a high-temperature channel, which is necessary for maximum light emission. The effect is emphasized in the oscillogram of figure 5*g*, which also includes a current pulse oscillogram. The discharge took place across an 0·5 cm. gap in hydrogen at 18 lb./in.^{2} abs., with a current of 120 amp. lasting for about 4·5 μsec. Figure 5*h* shows, for comparison, a discharge in nitrogen at 10 lb./in.^{2} abs., for the same current pulse.

The introduction of small amounts of oxygen into the argon has a similar effect to that caused by the introduction of like amounts of hydrogen. Figure 8*a* and *b* (plate 9) show, respectively, discharges across an 0·5 cm. gap in argon at 18 lb./in.^{2} abs., and in argon at the same pressure but containing 5 cm. Hg abs. of oxygen. The current pulse was 120 amp. lasting for 3 μsec. The addition of the oxygen causes an increased stability of the after-glow for each spark, but the after-glow still persists for the same length of time, and, as in the case of the addition of hydrogen to argon, there is no evidence of a marked 'quenching' effect on the duration of the emitted light, though it is known that such quantities of oxygen or hydrogen in argon have a noticeable effect on the electrical characteristics of the gas, and are effective in quenching metastable levels.

Discharges across a 0·5 cm. gap in helium at 45 lb./in.^{2} abs. and in a mixture of helium and argon in equal proportions at 18 lb./in.^{2} abs., produce the light pulses shown in figure 8*c* and *d* respectively. The current pulse, of 120 amp. for 4 μsec., is the same as that given in figure 5*f*. The introduction of helium into argon does not make any appreciable difference until the proportion of helium present is high, and even when the mixture contains equal amounts of argon and helium the after-glow time is still not less than that produced by 10 % of hydrogen in argon.

The rapidity of response of the light emission from a spark channel with sudden changes of current was investigated with an oscillatory current wave-form, as shown in figure 9*a*, plate 9, which also includes the corresponding light pulse for the discharge across an 0·5 cm. gap in oxygen at 7 lb./in.^{2} abs. The current amplitude for the first pulse of the oscillation was 120 amp. The light pulses shown in figure 9*b*–*e* (plate 9) refer respectively to discharges across a 0·5 cm. gap in hydrogen at 18 lb./in.^{2} abs., nitrogen at 10 lb./in.^{2} abs., argon at 19 lb./in.^{2} abs. and helium at 45 lb./in.^{2} abs. A fairly close relation between the light and current variations is obtained in hydrogen (short after-glow), but the transitions in light emission become less sharply defined for the other gases, especially for argon (long after-glow). These results are in accordance with expectations from the study of the single current pulse discharges. However, other interesting features of the argon discharge, for the oscillatory current wave, are the continued maintenance of light at a relatively high value, as compared with the other gases, and its steadiness for the series of sparks observed, in comparison with the variable nature of the after-glow when no current is flowing.

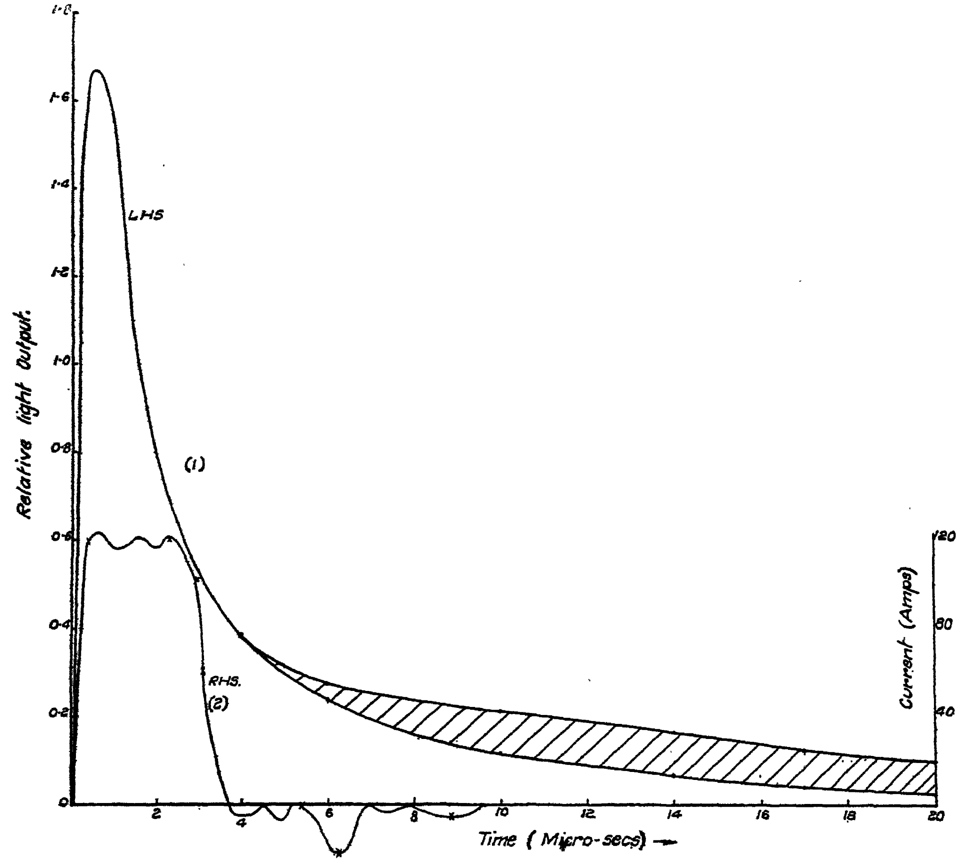


FIGURE 7*a*. Curves relating (1) light output from an argon spark as a function of time, (2) variation of discharge current with time.

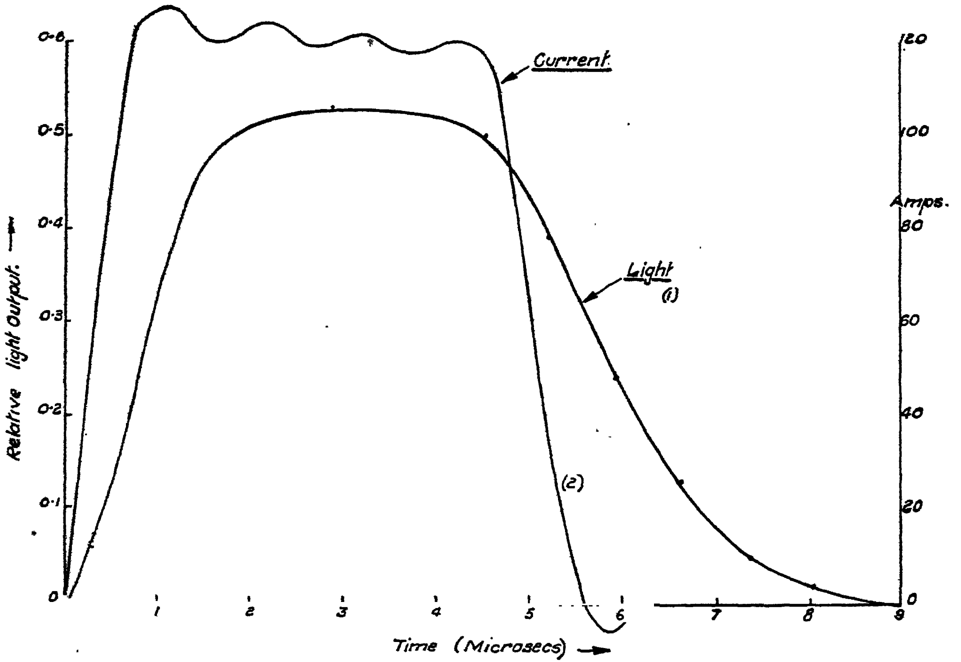


FIGURE 7*b*. Curves relating (1) light output from a hydrogen spark as a function of time, (2) variation of discharge current with time.

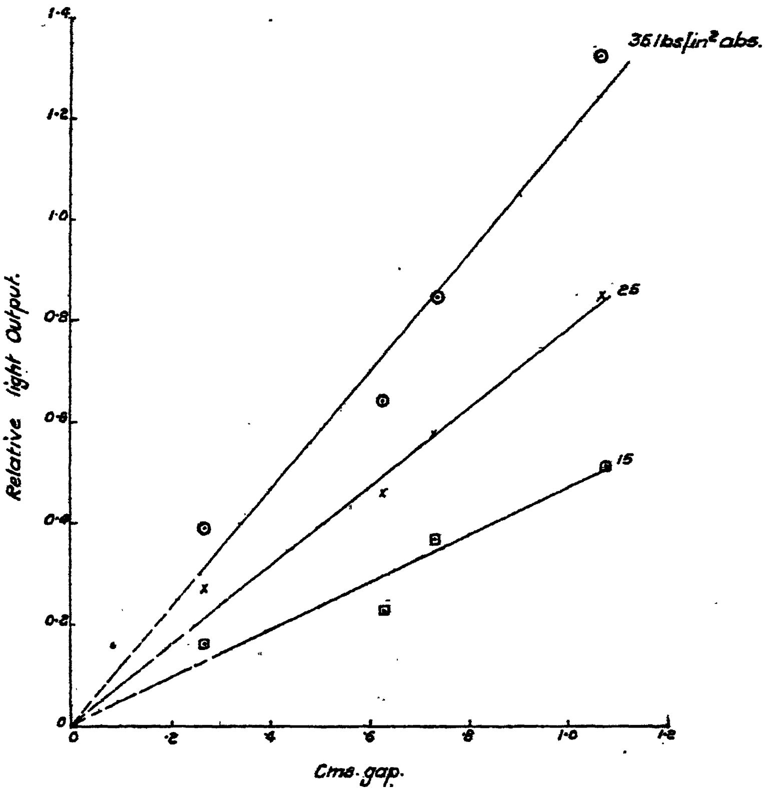


FIGURE 10*a*. Relation between maximum light output and gap length for spark discharges in argon at various pressures.

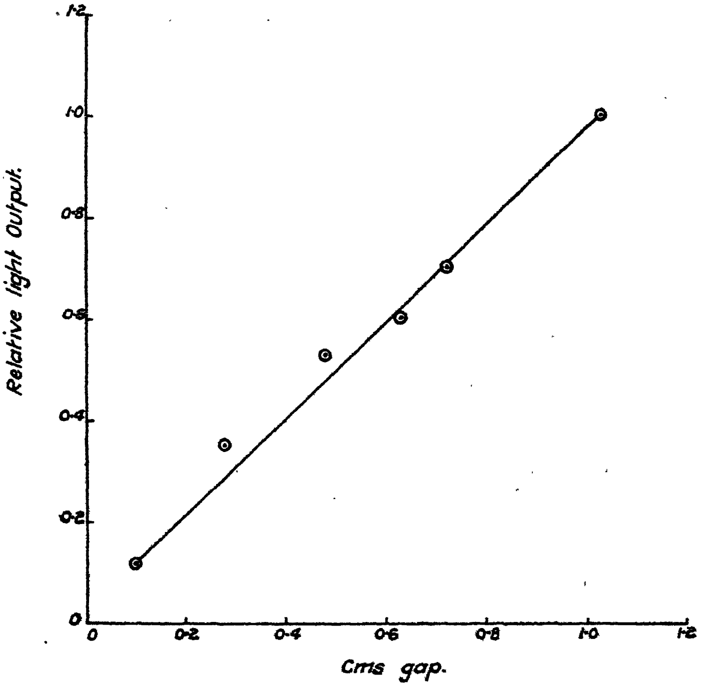


FIGURE 10*b*. Relation between maximum light output and gap length for a spark discharge in hydrogen at atmospheric pressure.

The peak-light emission varies directly as the gap length, in argon and hydrogen, for a given gas pressure and discharge current, as shown in figures 10*a* and *b*. This indicates that the channel is closely uniform throughout its length, except for the regions in the immediate vicinity of the electrodes. It is also in agreement with the records, given later in this section, which indicate a uniform voltage gradient along the spark channel.

A nearly linear relation also obtains for the variation of light with gas pressure (figure 11), for a given gap length and discharge current, over the range of gas pressures examined. The variation of peak-light emission with current is also a nearly linear one, for discharges in argon (figure 12*a*), in agreement with the results of Flowers (1943) for air. But in the case of hydrogen (figure 12*b*), while the light emission increases rapidly in a nearly linear manner, for currents up to about 150 amp., the rate of increase thereafter falls rapidly.

It is of interest to relate the light emission from the discharge with the rate of energy dissipation in the gap. The discharge tube was immersed in a heat-insulated oil-bath, and the discharge was operated with a square current pulse of 120 amp. and 3·6 μsec. duration, at a repetition rate of 200 pulses/sec. The average energy losses for a number of gap lengths in both argon and hydrogen were investigated, and considerations of the pulse duration and the repetition rate enabled the rate of energy expended during the pulse to be determined. Knowledge of the pulse-current amplitude then permits the calculation of the voltage drop across the gap, and the latter is plotted as a function of gap length in figure 13. For a given gap length the voltage drop is appreciably higher in hydrogen than in argon. Extrapolation of the curves for argon indicates that the cathode drop is about that to be expected in an arc, but a considerably higher value than the arc cathode drop is obtained for hydrogen. The voltage drop along the column of the discharges is roughly the same for both gases, at about 90 V/cm., and is appreciably above that corresponding to arc conditions on account of the constricted nature of the channel. Further investigations of the voltage drop by oscillographic methods confirm the calorimetric determination given above, and show that the voltage drop varies only slightly during the period of the pulse after falling from a high value during the first 0·1 μsec. Curves relating the light emission with the energy dissipation in the gap are given in figure 14. A nearly linear relation is obtained, though at the lower values the lines must curve and tend towards the origin. The linear variation of light output with current (figure 12*a*) is now explicable since the data of figure 14 show light output to be proportional to the wattage dissipated in the gap, and other experiments (D. E. M. Garfitt, unpublished) have shown that the voltage drop in the gap was sensibly independent of current over the range covered (up to 40 amp.).

With a given gas, for the range of gas pressures and currents used, it has been considered sufficiently accurate to assume a relatively constant spectral distribution, and to plot the relative changes in total light emission, as recorded by the multipliers, with no corrections on this account. Direct comparisons between the light emitted from discharges from different gases cannot be made on this basis without regard to

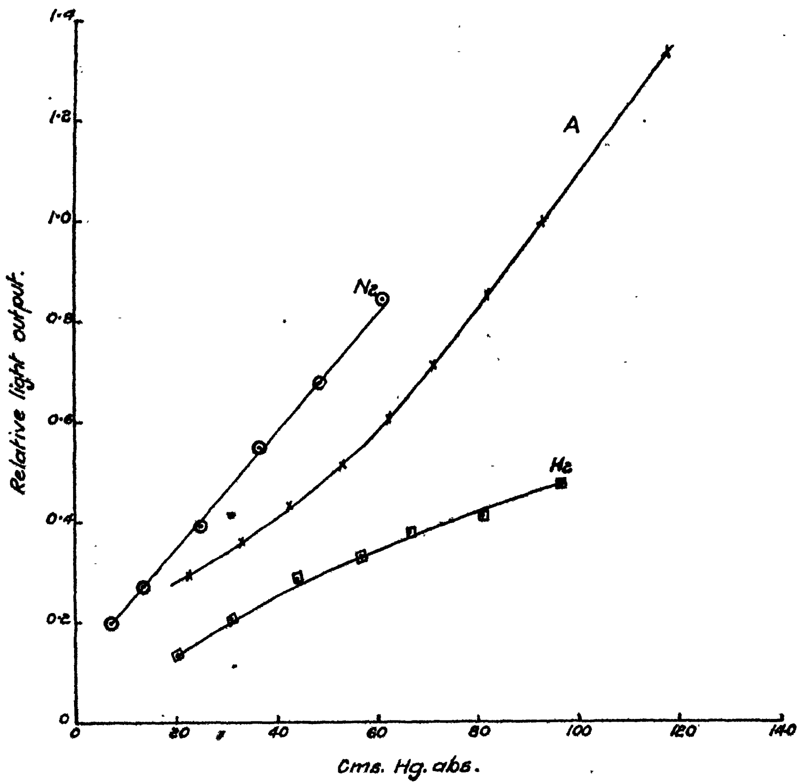


FIGURE 11*a*. Relation between maximum light output and gas pressure for a spark of length 0·63 cm. in argon, nitrogen and hydrogen.

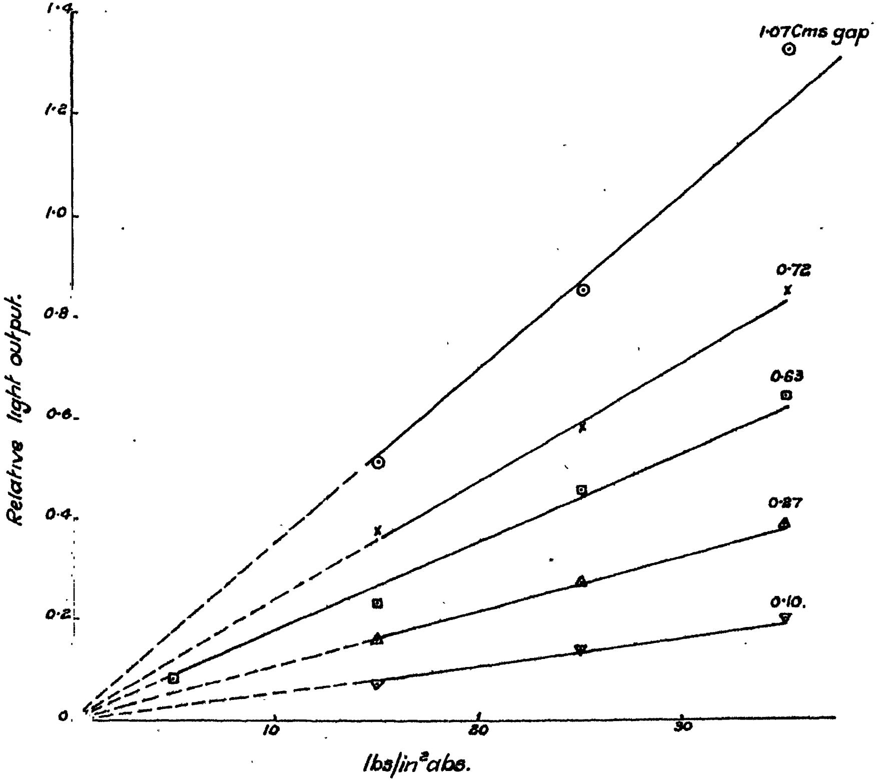


FIGURE 11*b*. Relation between maximum light output and gas pressure for sparks of various lengths in argon.

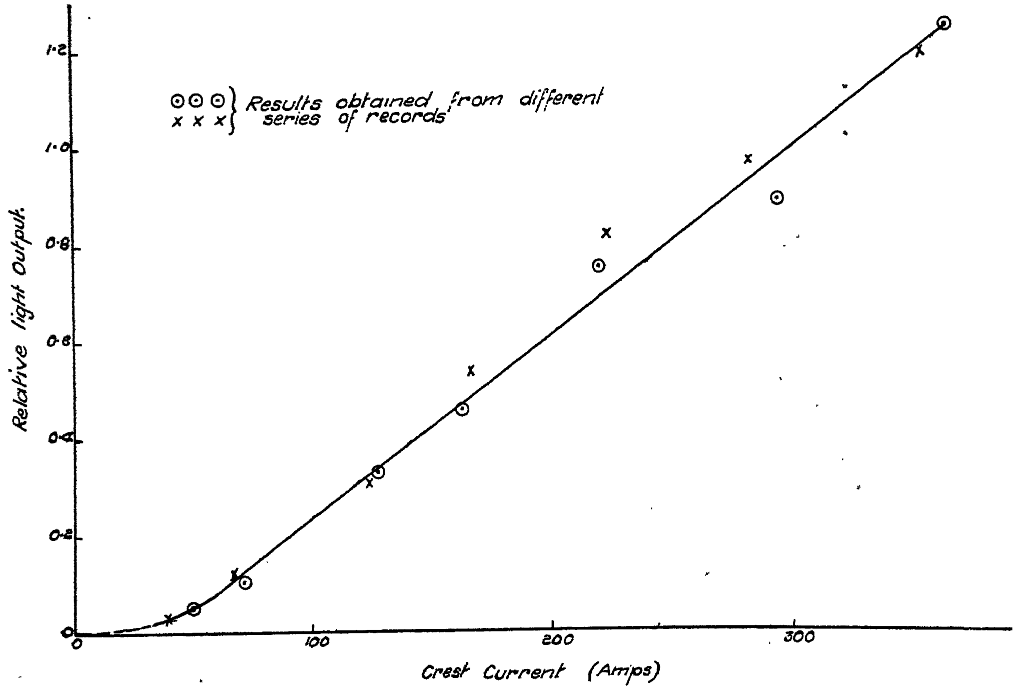


FIGURE 12*a*. Curve relating the maximum rate of light output from a spark discharge between tungsten electrodes separated by 0·5 cm. in argon 19 lb. per in.^{2} abs. as a function of the maximum discharge current.

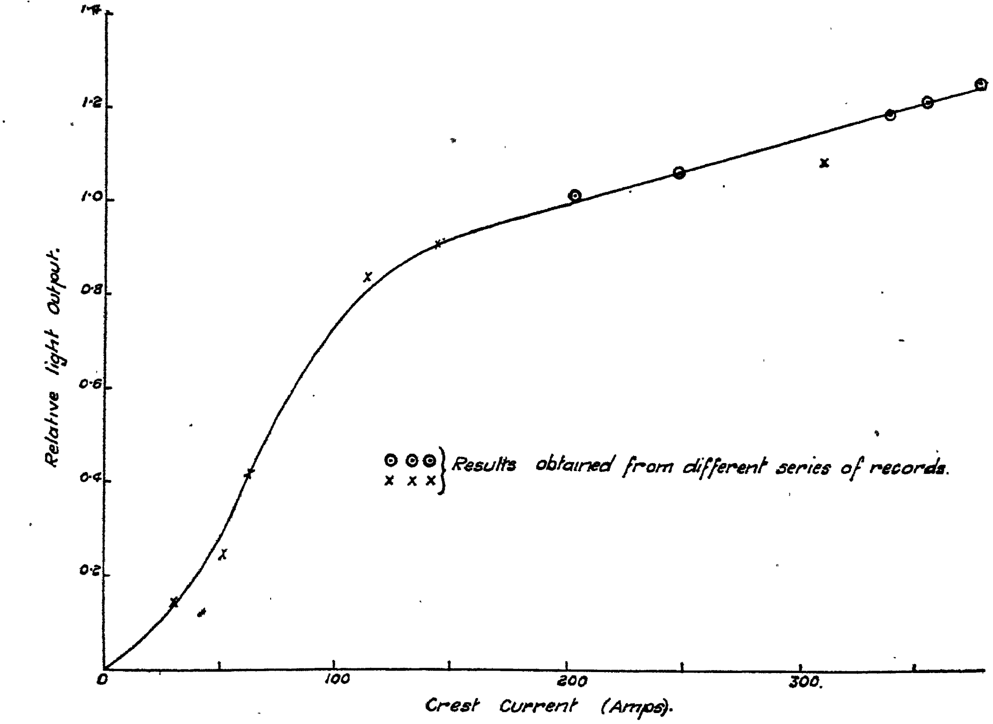


FIGURE 12*b*. Curve relating maximum light output from a spark discharge between tungsten electrodes separated by 0·5 cm. in hydrogen at 17 lb./in.^{2} abs. as a function of the maximum discharge current.

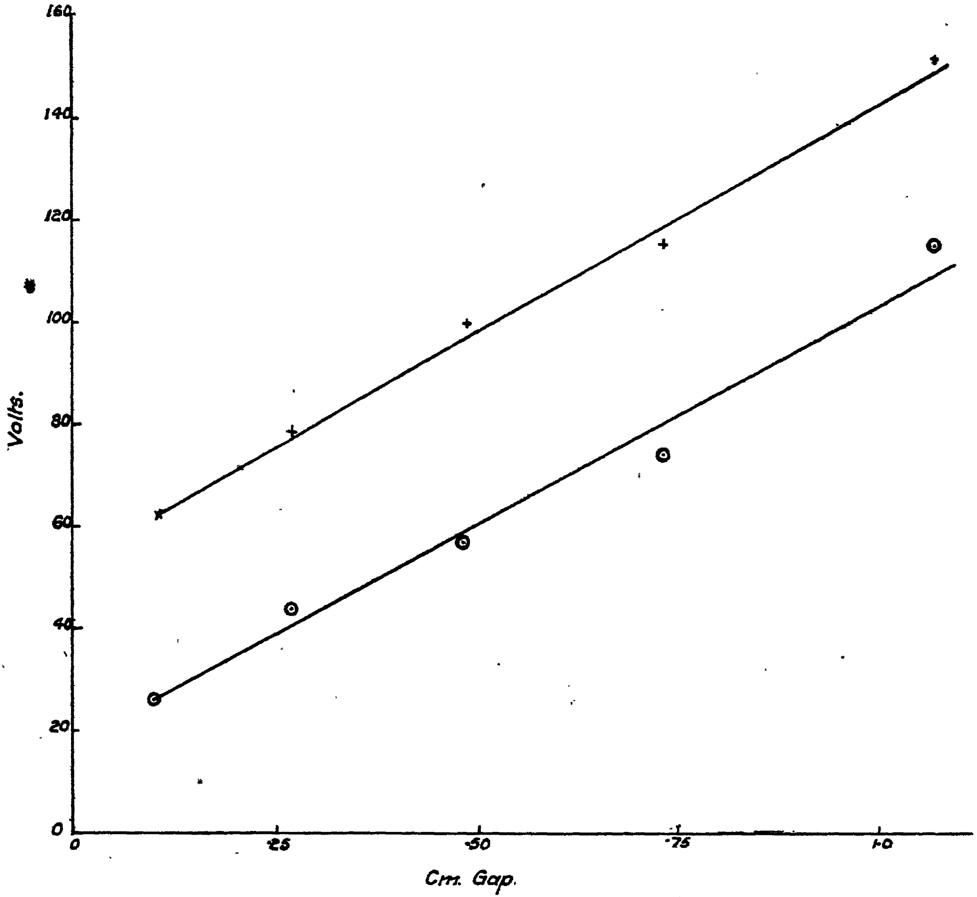


FIGURE 13. Variation of voltage drop across a spark discharge as a function of gap length for a current of 120 amps. + Hydrogen at 16 lb./in.^{2} abs. ⊙ Argon at 16 lb./in.^{2} abs.

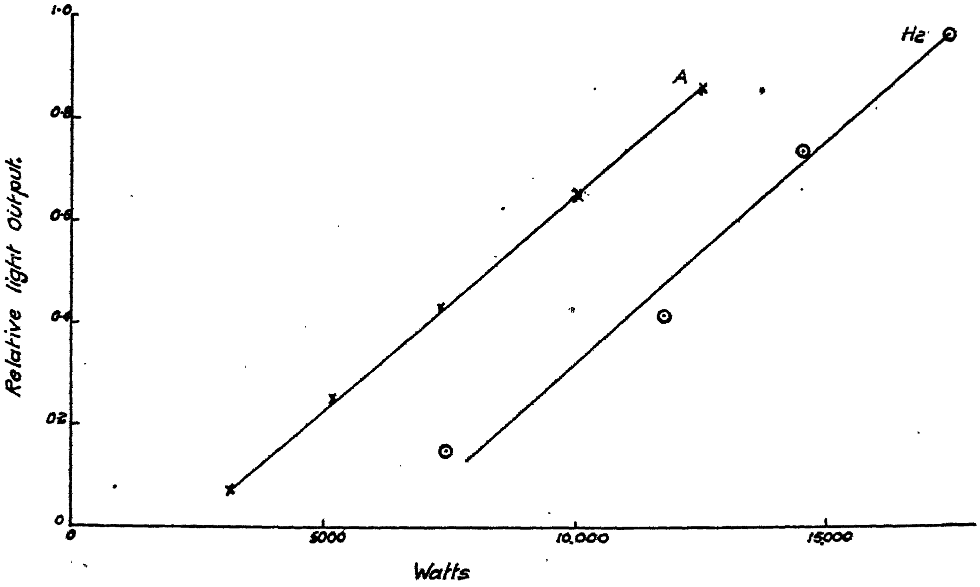


FIGURE 14. Relation between maximum light output for sparks in argon and hydrogen at atmospheric pressure and rate of energy dissipation.

the differences in spectral distribution, and no attempt is made here to compare the light on an absolute scale. Experiments were made, however, with various filters, to obtain an indication of the spectral changes in the light emitted during the discharges as a function of time. The filters were introduced between the spark and the window of the multiplier, and some of the resultant oscillograms for various discharges are shown in figure 15 (plate 10). The current pulse has an amplitude of 120 amp.; other details of the discharges are given in table 3.

TABLE 3

figure	gas	pressure lb./in.^{2} abs.	gap length cm.	pulse length μsec.	Wratten filters used
15*a*	A	19	0·5	4	none, 1, 4, 15
b	H_2	18	0·5	4	none, 2*a*, 15, 27
c	A/H_2	11·5/8·5	1·0	3	none, 1, 2*a*, 4, 15, 27

The effect of the various filters (§2 and figure 3), which successively cut off the light of lower wave-length reaching the multiplier, is most pronounced in the reduction of the amplitude of the light which occurs during the initial stages of the discharge, and indicates the presence of more intense ultra-violet emission at this stage in the spark development than in the later stages. This is to be expected on account of the high fields and ion concentration which occur during the initiation of the spark channel. The effect is most noticeable in the reduction of the peak light for the argon and argon-hydrogen discharges. The abnormally high light output at the commencement of the current pulse in argon (figure 5*a*), as compared with hydrogen (figure 5*c*) where the light output reaches a maximum slowly and only falls at the end of the current pulse, appears to be connected with the existence of a restricted channel in the early part of the discharge after complete breakdown of the gas has occurred.

Some preliminary experiments were made to examine the effect of electrode materials on the nature of the light emitted. For this purpose a discharge across a 0·3 cm. gap in air was observed, with a square pulse of amplitude 120 amp., and some of the records so obtained are shown in figure 16*a*, *b*, *c* (plate 10) for electrodes of tungsten, copper and magnesium respectively. Figure 16*d* shows the current oscillogram. The light emitted during the initial stages of the discharge is not appreciably affected by the changes in electrode material, as would be expected, since the light emission is governed here by the characteristics of the gas. During the later stages, when metal vapour has had time to diffuse into the gap, considerable differences become evident. The after-glow is particularly prolonged for magnesium electrodes and is least for tungsten. This can probably be explained not only by differences in spectral emission for the metals used, but also by the differences in the total amount of metal vaporized, which will be the greater for magnesium, with a lower boiling point.

4. Discussion of results

4·1. *Hydrogen discharge*

The light output during and after the passage of a square current pulse in hydrogen is shown in figure 5*e–g* (plate 7).

The after-glow which may be caused by thermal excitation in the hot spark channel, or by recombination, is much longer than the mean life (10^{-8} sec.) of the excited states (Rayleigh 1944), and to assess the importance of these processes a knowledge of the channel temperature is necessary. The Stark broadening (figure 17*a*, plate 11) observed with the 120 amp. sparks at 1 atm. pressure indicates an inter-ionic field of about 330×10^3 V/cm. (von Traubenburg *et al.* 1930) and thus an ion concentration of about $3{\cdot}5\times10^{18}$/c.c. or 13 % ionization (Margenau & Watson 1936). This result is clearly only an approximation and represents an upper limit for the estimate of ion concentration.

In a time of 1 μsec. it may be considered that thermal equilibrium is virtually established in the spark channel, as during this time many thousands of electronic collisions have occurred. In arcs, at much lower temperatures, times of about 1 msec. are required to establish thermal equilibrium (Witte 1934).

On the assumption of thermal equilibrium, the following form of Saha's equation enables spark channel temperatures to be calculated from a knowledge of ion concentrations:

$$\log_{10}\frac{Px^2}{1-x^2}=\frac{-5050\,V_i}{T}+2{\cdot}5\log_{10}T-6{\cdot}5, \qquad (1)$$

where P = pressure in atmospheres, T = temperature (° K), x = degree of ionization and V_i = ionization potential (13·5 V for ionization of H_1 and 18 V for ionization to H_1^+ of H_2). Molecular dissociation (Langmuir 1915; Finkelnburg 1931) will be almost complete in the channel. Calculations have been made only for V_i of 15 (V_i for argon = 15·7 V) and 18 V, because (Unsöld 1938) the high inter-atomic fields may affect the ionization potential, thus making an accurate estimate of V_i difficult. The effect of initial vapour on V_i is probably negligible, as the spark spectra indicate; figure 17*a* and *b* (plate 11) respectively show hydrogen discharges between electrodes of copper/tungsten alloy for a 3 μsec. pulse and a 5 mm. gap and a 1 μsec. pulse with a 12 mm. gap.

Table 4 gives results of calculations from equation (1). N_i = number of electrons or positive ions/c.c. in the channel. In the presence of an electric field N_e (electrons) will differ from N_i, and a modified form of equation (1) has then to be used. As the after-glow conditions refer to zero field, it is felt justifiable at least as a first approximation to put $N_e \doteq N_i$. Table 4 gives several pressures, since the true channel pressure varies with its temperature (constant-density conditions obtain) so that, for example, P = 20 atm., $T \equiv 6000°$ K, and P = 50 atm., $T \equiv 15{,}000°$ K. Molecular dissociation also affects the pressure values.

Now if A is the channel cross-sectional area, e and v the electronic charge and drift velocity respectively, then

$$I = N_i A e v. \tag{2}$$

With the current of 120 amp. used in the majority of the experiments the channel radius of the discharge in hydrogen appears to be about 0·75 mm. so that

$$v = \frac{4{\cdot}2 \times 10^{22}}{N_i}. \tag{3}$$

TABLE 4

$T°$ K	$P = 1$ atm.	$P = 20$ atm.	$P = 50$ atm.	v cm./sec.	X V/cm.
		$V_i = 15$ V, N_i ions/c.c.			
10,000	$2{\cdot}5 \times 10^{17}$	$5{\cdot}6 \times 10^{16}$	$3{\cdot}5 \times 10^{16}$	$1{\cdot}0 \times 10^{6}$ ($P = 34$ atm.)	200
12,000	$1{\cdot}3 \times 10^{18}$	$2{\cdot}8 \times 10^{17}$	$1{\cdot}8 \times 10^{17}$	$2{\cdot}0 \times 10^{5}$ ($P = 41$ atm.)	40
15,000	$7{\cdot}0 \times 10^{18}$	$1{\cdot}6 \times 10^{18}$	$1{\cdot}0 \times 10^{18}$	$4{\cdot}2 \times 10^{4}$ ($P = 51$ atm.)	8
		$V_i = 18$ V, N_i ions/c.c.			
10,000	$4{\cdot}3 \times 10^{16}$	$9{\cdot}6 \times 10^{15}$	$6{\cdot}1 \times 10^{15}$	$5{\cdot}7 \times 10^{6}$ ($P = 34$ atm.)	1,100
12,000	$3{\cdot}1 \times 10^{17}$	$6{\cdot}7 \times 10^{16}$	$4{\cdot}1 \times 10^{16}$	$9{\cdot}1 \times 10^{5}$ ($P = 41$ atm.)	180
15,000	$2{\cdot}3 \times 10^{18}$	$1{\cdot}6 \times 10^{18}$	$3{\cdot}4 \times 10^{17}$	$1{\cdot}2 \times 10^{5}$ ($P = 51$ atm.)	24

In table 4 values of v as determined from equation (3) are given corresponding to the various temperatures on the assumption of constant gas density in the channel. From these values it is possible to estimate X, the field strength along the channel, by considerations of electron mobility (Loeb 1939). Agreement with the experimentally determined value for X, of about 90 V/cm. (figure 13), is obtained with $V_i = 18$ V at a channel temperature between 12,000 and 15,000° K, when the ion density is approximately 10^{17} ions/c.c.

The energy input to the channel, assumed to have a radius of 0·75 mm. and a length of 1 cm., is 18 kW at 120 amp. at the end of the current pulse. The energy is emitted partly as a continuum, but largely as a line spectrum (figure 17 *a* and *b*).* If the heat capacity and rate of heat loss of the channel were known, the cooling time could be calculated for a range of 12000–8000° K (equivalent to $N_{i_0} = 10^{17}$ and $N_i = 5 \times 10^{15}$, see below). Using the results obtained above, assuming a reasonable value of emissivity from line breadths, etc., and taking $C_v = 17{\cdot}5$ cal./g.mol.

---

* *Note added in proof:* Energy dissipated in other ways, i.e. by gas movement etc. is probably negligible in the very short times involved here (Flowers 1943).

because of dissociation (Pier 1909; Langmuir 1915), the cooling time is found to be of the correct order, i.e. about 5 μsec.

It is necessary now to consider the recombination process in more detail. The recombination coefficient α is defined by

$$dN_{i_0} = -\alpha N_{i_0}^2 \, dt, \tag{4}$$

where N_{i_0} and N_i are the values of ion or electron concentration, assumed equal, at times $t = 0$ and $t = t$ respectively. If α is independent of N_i then

$$N_i = \frac{N_{i_0}}{1 + N_{i_0}\alpha t}. \tag{5}$$

There do not appear to be any data on electron/ion recombination at these very high temperatures and ion concentrations. Mohler (1937) found $\alpha = 10^{-10}$ for hydrogen in a particular case of a low-pressure discharge, but this value should be appreciably higher than that obtained in sparks. Indeed, the value $\alpha = 10^{-11}$ substituted in equation (5) seemed to give the best agreement with the previous Stark effect and Saha theory calculations, and gave the calculated data of table 5 ($t = 2$ μsec.).

TABLE 5

N_{i_0}	N_i	N_{i_0}/N_i
10^{20}	5×10^{16}	2×10^{3}
10^{18}	$4{\cdot}7 \times 10^{16}$	21
10^{17}	$3{\cdot}3 \times 10^{16}$	3
10^{16}	$8{\cdot}3 \times 10^{15}$	1·2
10^{15}	$\sim 10^{15}$	~ 1
10^{10}	$\sim 10^{10}$	~ 1

It is seen, e.g. from figure 6*g*, that the light intensity has fallen to about 1/10 in 2 μsec., so that if recombination is responsible for the after-glow

$$N_{i_0}/N_i = \sqrt{10}, \quad \text{and therefore} \quad N_{i_0} \sim 10^{17} \text{ from table 5.}$$

If α is lower than 10^{-11}, then the recombination process possibly becomes insignificant in comparison with thermal excitation. It has been assumed for simplicity that no fresh ions are produced during the cooling period, which cannot be true. To give the observed rate of decrease of light intensity α should then be greater than 10^{-11}.

The number of recombination processes per sec. in a channel of length 1 cm. and radius 0·75 mm. is $1{\cdot}8 \times 10^{21}$ if $N_{i_0} = 10^{17}$ ions/c.c. and $\alpha = 10^{-11}$. It has been shown (§ 2) that to give a full screen deflexion on the oscillograph a spark at 15 cm. from the multiplier cathode must emit $3{\cdot}6 \times 10^{19}$ quanta/sec. for $\lambda = 4861$ A. This number of quanta is much less than the number ($1{\cdot}8 \times 10^{21}$) of recombinations and could, it is reasonable to suppose, be supplied at $\lambda = 4861$ A. Hence the recombination process alone could be responsible for the after-glow in hydrogen.

The relative populations of two states, energies ϵ_1 and ϵ_0, of which ϵ_0 may be taken as the ground state, are given by

$$\frac{N_i}{N_0} = \frac{W_1 e^{-\epsilon_1/\mathbf{k}T}}{W_0 e^{-\epsilon_0/\mathbf{k}T}} \tag{6}$$

for thermal equilibrium, where **k** and T are respectively Boltzmann's constant and the absolute temperature. W_1 and W_0 are the statistical weights. For hydrogen $\epsilon_1 \sim 12\,\mathrm{eV} \equiv 1{\cdot}9 \times 10^{-11}$ ergs and $W_1 \sim W_0$, so that $N_i/N_0 \sim 9 \times 10^{-5}$ at 15,000° K. N_0 is $2{\cdot}7 \times 10^{19}$ atoms/c.c. at N.T.P. or in the spark channel for constant density conditions, and the mean life of the excited atoms is about 10^{-8} sec., so that about 4×10^{21} quanta/sec. are emitted for the whole spark ($r = 0{\cdot}75$ mm., $l = 1$ cm.). This quantity of light, which covers a wide spectral range, also is greatly in excess of that required at $\lambda = 4861$ A to give the observed multiplier output.

In conclusion it seems that either thermal excitation or recombination could account for the after-glow. It is not possible from the present preliminary work to decide which process predominates, but spectrographic and other methods of observation are being pursued. The recombination continuum on the short wavelength side of the Balmer series limit has been observed. It appears that the persistence of energetic electrons from the discharge, in the after-glow period, may be ignored; the mean electron energy falls to a negligible amount in about $0{\cdot}1\,\mu$sec. for both hydrogen and argon sparks.

4·2. *Argon discharge*

Most of the arguments advanced in § 4·1 refer qualitatively to spark channels in argon when, however, dissociation complications are absent.

It seems (figures 5*a*, 5*g*) that argon spark channels are established more rapidly, in that the rate of light emission increases more rapidly. This would be expected from the relative values of electronic mean free paths and energy transfers at collisions (Healey & Reed 1941). The heat capacity of an argon channel is lower than for hydrogen, because C_v (argon) is about one-sixth of C_v (hydrogen) at 10,000° K, and so will more rapidly reach a given temperature for the same input energy. Unfortunately the strong argon after-glow masks the effect of channel cooling.

Argon (figure 17*c*), helium (figure 17*d*) and neon (figure 17*e*) spectra are shown for 30 amp. and 55 lb./sq.in. abs. pressure.

It is likely that the after-glows in the rare gases, particularly argon, are due to metastable atoms. The estimation of such effects is difficult at present, as extrapolation from low pressure results appears unjustified.

The visible after-glow in argon, which is that recorded by the glass bulb system, is not greatly affected as to time duration by the addition of quenching agents, such as hydrogen or oxygen in moderate amounts (figures 5*a*, *b* and 8*a*, *b*), which exert a marked influence in low pressure discharges. The presence of quenching agents removes the variations in after-glow as observed for consecutive sparks. Such variations are probably connected with the fact that the after-glow is often (figure 6*n*) though not always (figure 6*m*) concentrated into discrete regions. A similar effect is noticed with moving camera records of long sparks and lightning where interpretation is further complicated by current flow. It is possible that the localized regions of after-glow can be associated with differences in space charge accumulation in the discharge, as produced by variations of voltage gradient along the channel. A further investigation of this phenomenon is planned.

The variation with current of the light emission from an argon spark can be explained on the basis of recombination in the following approximate manner. From equation (4) the light emission L per unit length of channel of radius r is given by

$$L \propto \frac{dN_i}{dt} \propto N_i^2 r^2.$$

It has been shown in equation (2) that $N_i \propto I/r^2 v$ and therefore

$$L \propto \frac{I^2}{r^2 v^2}.$$

Now Flowers (1943) has shown experimentally that I/r^2 is constant for established spark channels. Also, considerations of the variation of electron mobility with temperature and gas pressure indicate that the electron drift velocity v is roughly proportional to the voltage gradient along the channel, on the assumption of constant gas density, and in unpublished work by D. E. M. Garfitt it has been shown that the voltage drop across spark gaps in argon varies little over the range of parameters used in the present investigation. If then we consider both I/r^2 and v to be constant, expression for L becomes

$$L \propto I,$$

in agreement with the present observations (figure 12*a*). Alternatively, if the light emission is caused by thermal excitation, then

$$L \propto N_i r^2,$$

or on substitution from equation (2)

$$L \propto I/v.$$

Again, for the reasons discussed in the preceding paragraph, we may consider v to be roughly constant, so that

$$L \propto I.$$

5. CONCLUSIONS

Observations of the variations with time of the light emission from spark discharges indicate that the after-glow in hydrogen is probably a thermal effect, whilst in argon it is due largely to other causes, of which the persistence of atoms in metastable states is the most likely.

From considerations of the spectra of the discharge and the energy balance in the channel it is deduced that the density of ionization is about 10^{17} ions/c.c. for the sparks observed. Channel temperatures, as determined on the basis of Saha's equation, are shown to be about 10,000–15,000° K.

The mechanism of light emission from the channel can be explained in terms either of thermal excitation or of electron/ion recombination. Further investigations which are being undertaken are necessary to decide the relative importance of these two processes.

The authors wish to express their thanks to Mr L. J. Davies and Mr H. R. Ruff for the supply of the calibrated projector lamp and filters which were used in the calibration of the multiplier, and to Mr H. Warren, Director of Research, British Thomson-Houston Co. Ltd., for making this cooperation possible.

The stimulating discussions which the authors have had with Professor K. G. Emeléus have greatly contributed to the research work. Professor Emeléus, Dr F. Llewellyn Jones and Dr H. Fröhlich have given their comments on the original manuscript of this paper, and their assistance is gratefully acknowledged. Thanks are due to Dr T. E. Allibone and Mr F. R. Perry for their continued interest and encouragement, and to Sir Arthur P. M. Fleming, C.B.E., D.Eng., Director of Research and Education, Metropolitan-Vickers Electrical Co. Ltd., for permission to publish this paper.

REFERENCES

Finkelnburg, W. 1931 *Z. Phys.* **70**, 375.
Flowers, J. W. 1943 *Phys. Rev.* **64**, 225.
Healey, R. H. & Reed, J. W. 1941 *The behaviour of slow electrons in gases.* Sydney: Amalgamated Wireless, Australasia.
Kennelly, A. E. 1917 *Artificial electric lines.* New York: McGraw-Hill.
Langmuir, I. 1915 *J. Amer. Chem Soc.* **37**, 417.
Loeb, L. B. 1939 *Fundamental processes of electrical discharge in gases.* New York: McGraw-Hill.
Loeb, L. B. & Meek, J. M. 1940 *J. Appl. Phys.* **11**, 438.
Margenau, H. & Watson, W. W. 1936 *Rev. Mod. Phys.* **8**, 22.
Mohler, F. L. 1937 *Bur. Stand. J. Res.* **19**, 559.
Pier, M. 1909 *Z. Elektrochem.* **15**, 536.
Rayleigh, Lord 1944 *Proc. Roy. Soc.* A, **183**, 26.
Traubenburg, H. R. von, Gebauer, R. & Lewin, G. 1930 *Naturwissenschaften*, **18**, 329.
Unsöld, A. 1938 *Ann. Phys., Lpz.*, **33**, 607.
Witte, H. 1934 *Z. Phys.* **88**, 415.

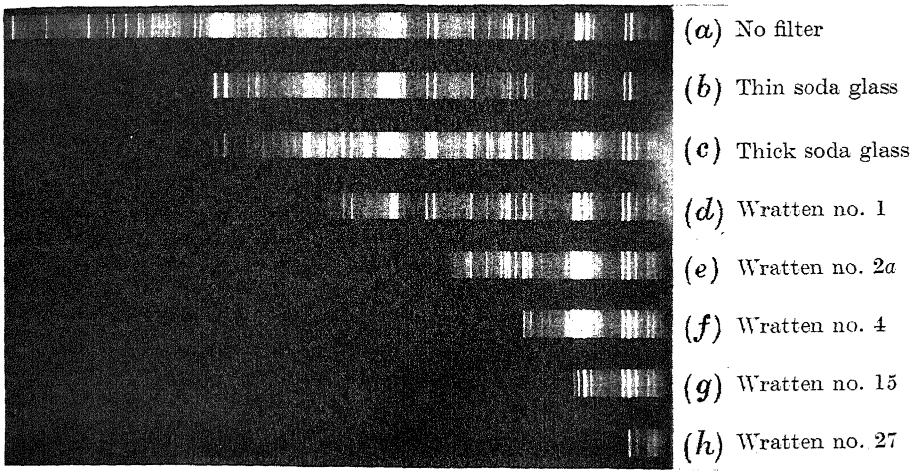
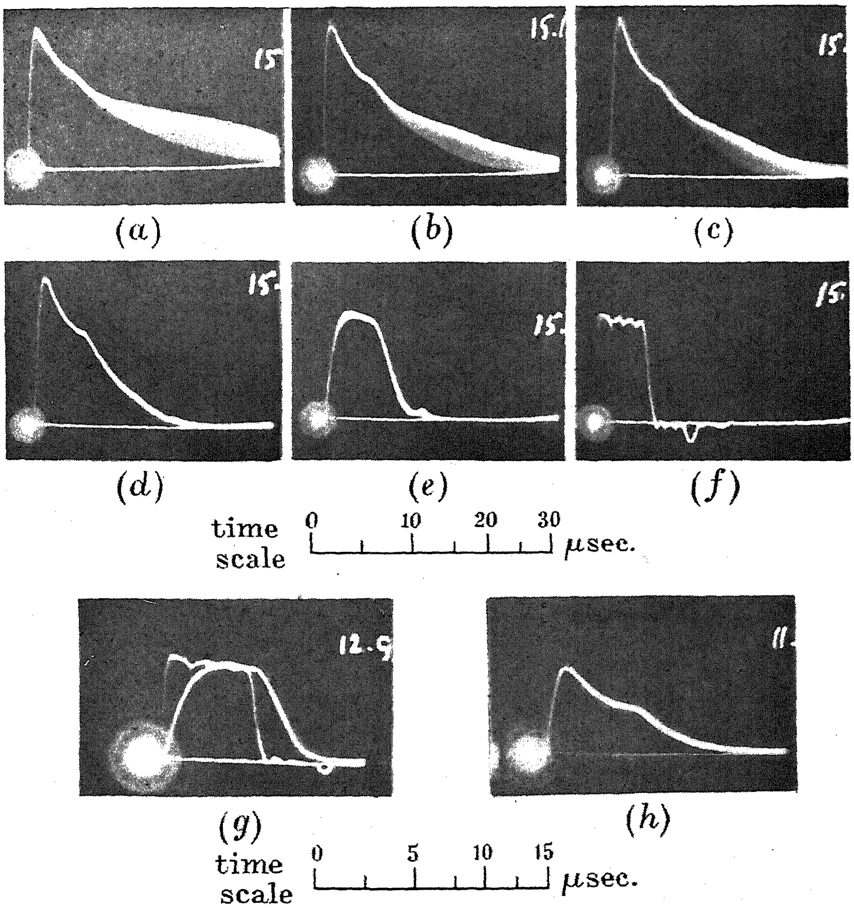


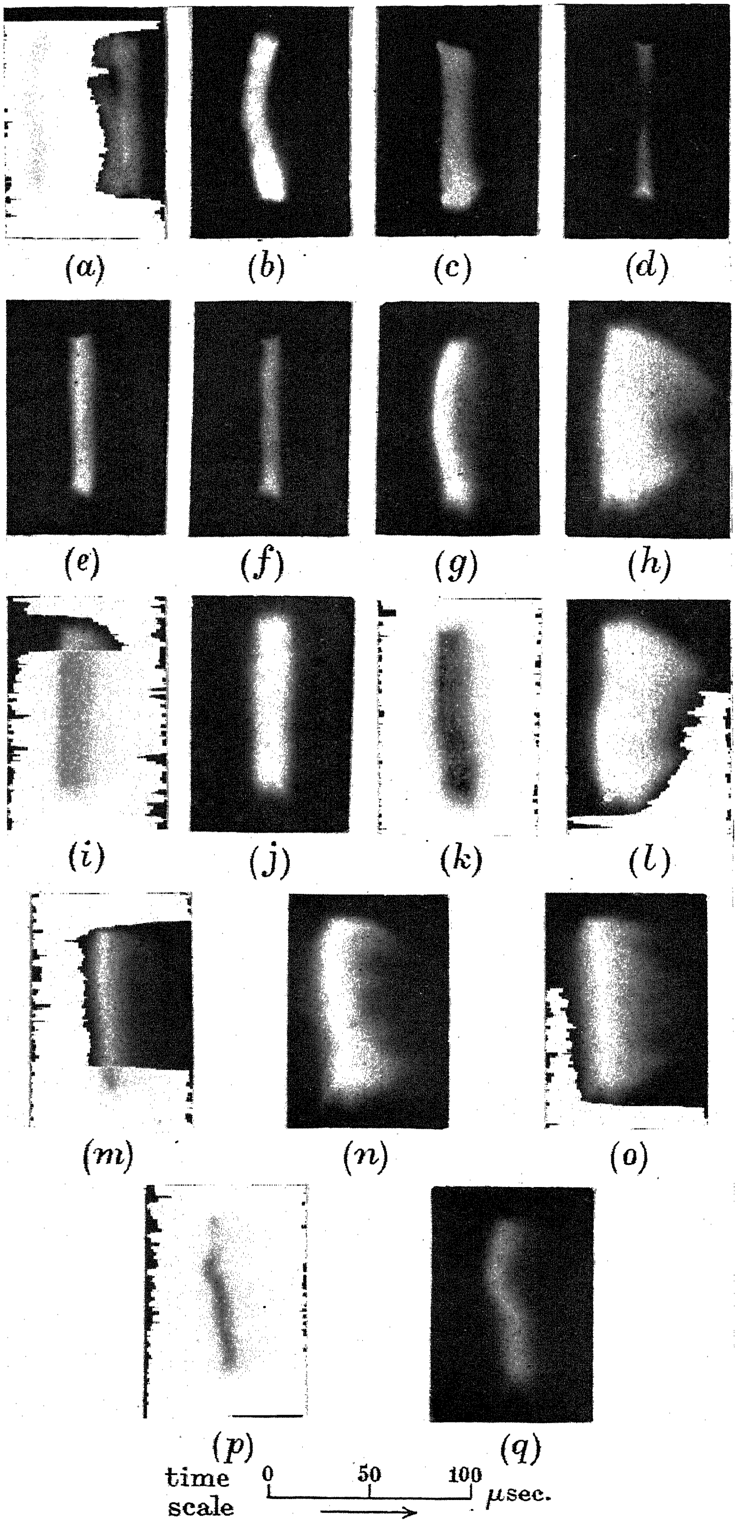
FIGURE 3



(*a*) (*b*) (*c*) (*d*) (*e*) (*f*) (*g*) (*h*)

FIGURE 5

The calibration of ordinates is the same for all light oscillograms.
To obtain relative light values, refer to figure 1, §§ 2, 3.


(*a*) (*b*) (*c*) (*d*)

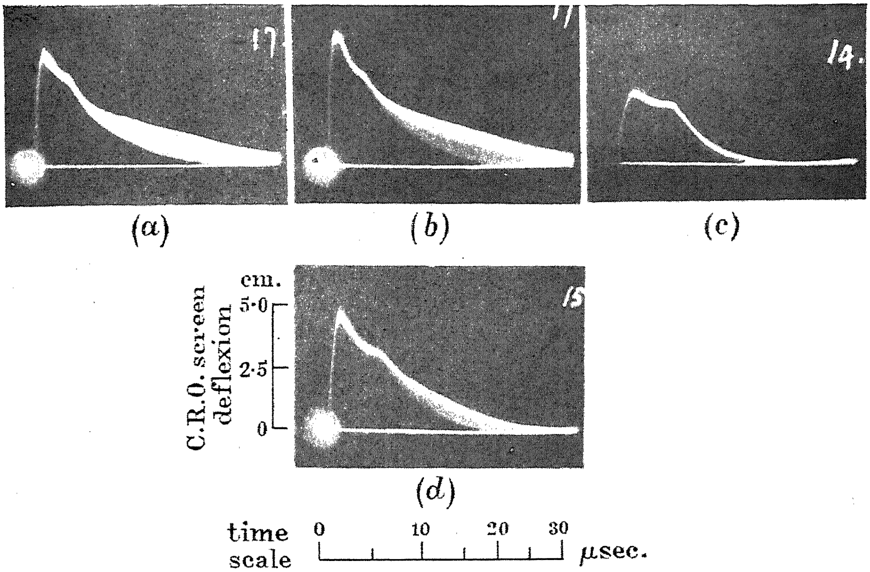
(*e*) (*f*) (*g*) (*h*)

(*i*) (*j*) (*k*) (*l*)

(*m*) (*n*) (*o*)

(*p*) (*q*)

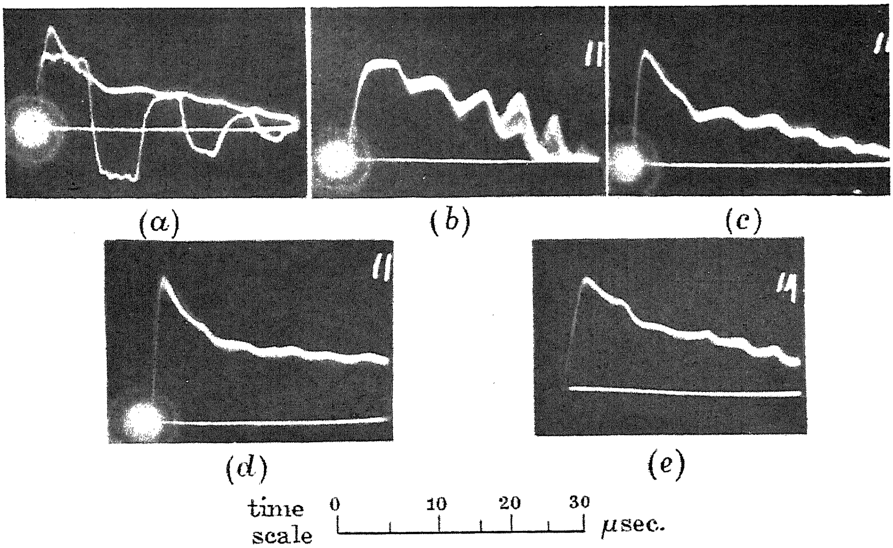
FIGURE 6



(*a*) (*b*) (*c*)

(*d*)

FIGURE 8



(*a*) (*b*) (*c*)

(*d*) (*e*)

FIGURE 9

The calibration of ordinates is the same for all light oscillograms. To obtain relative light outputs, refer to figure 1, §§ 2, 3.

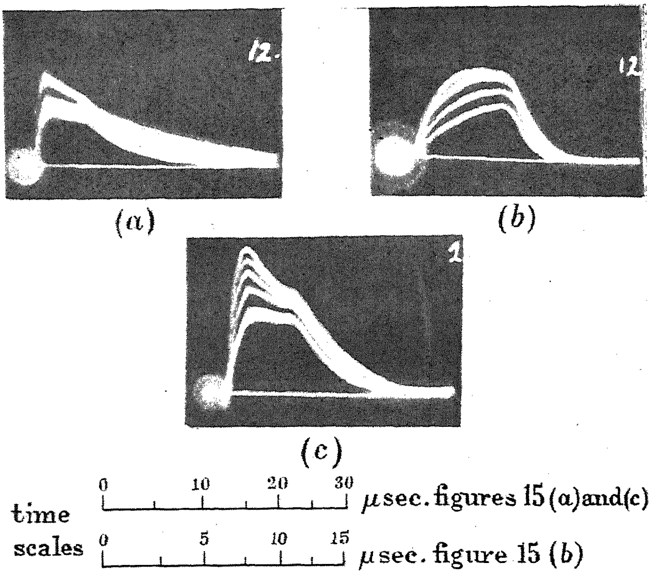


FIGURE 15

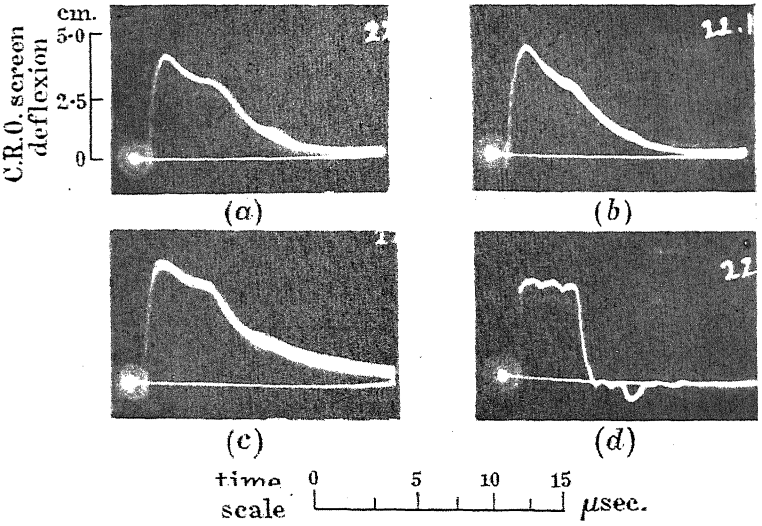


FIGURE 16

The calibration of ordinates is the same for all light oscillograms. To obtain relative light outputs, refer to figure 1, §§ 2, 3.

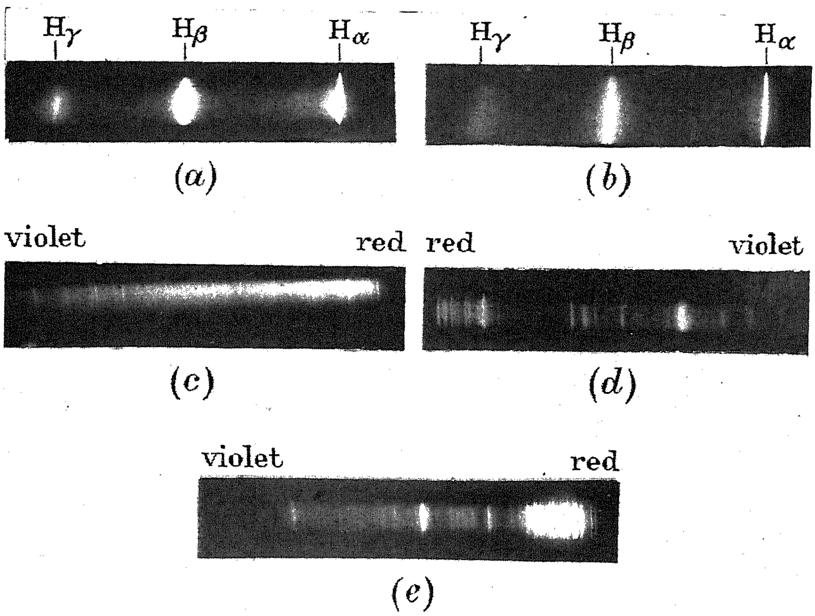


FIGURE 17

Further interferometric studies with mica
New multiple-beam fringes and their application

By S. Tolansky, Ph.D., *Manchester University*

(*Communicated by Sir Lawrence Bragg, F.R.S.—Received* 14 *March* 1945)

[Plates 12–14]

A description is given of multiple-beam interference precision methods for studying the surface topography and internal structural features of thin slips of mica. Three methods are described involving (*a*) new white light multiple-beam interference fringes, called here 'fringes of equal chromatic order', (*b*) an adaptation of monochromatic multiple-beam Fizeau fringes, (*c*) monochromatic multiple-beam non-localized fringes of high dispersion. The properties of the fringes of equal chromatic order are discussed and their applications reviewed. They are shown to be a powerful weapon for investigating interferometric properties of thin films. The directions of surface contours and of cleavage steps can be determined. Surface angles of only 0·003 min. of arc can be measured, over a length of 1 cm., this being at least 50 times better than the Rayleigh limit of the goniometer. The fringes reveal information about surface contours, local variations in chemical composition local occlusions, growth nuclei, birefringence, nature of cleavage surface, etc. The precision is such that the molecular lattice spacing of mica normal to the cleavage direction can be evaluated. The birefringence can be determined even if it is as low as 0·0001. Changes in thickness of a mica sheet are shown to be discontinuous.

The non-localized fringes of high dispersion also reveal cleavage steps and birefringence.

Introduction

A description is given here of three multiple-beam interference methods, used in this particular instance for the study of structural and surface topographical features of thin slips of mica. This account will deal with the experimental techniques employed, detailed analysis of specific mica features being communicated elsewhere. The procedures described have wide potential applications and can be employed with any type of approximately plane surface, and to any thin solid or liquid transparent medium which does not attack silver.

The three methods used involve (*a*) a new type of multiple-beam white light fringes, which for reasons developed fully elsewhere are called '*fringes of equal chromatic order*', (*b*) an adaptation of multiple-beam Fizeau fringes, using monochromatic light, (*c*) the multiple-beam non-localized monochromatic ring fringes of high dispersion, recently described by the writer. These three methods, each sensitive, together form a powerful optical combination, complementing each other; method (*b*) gives information about extended areas, method (*a*) refers to a selected line section, and method (*c*) supplies restricted information about small areas (crude points), and in some cases extended areas.

Of the three methods, which are independent, the most powerful is method (*a*), employing the multiple-beam white light fringes. This procedure solves, by inspection alone, problems intractable to other methods. It has wide possible application to thin film interferometry (low order of interference) and will be described first.

In earlier publications (Tolansky 1943*a*, 1944*a*, 1945) a description has been given of a *monochromatic* multiple-beam interference procedure for studying the topographical features of the surfaces of fairly flat crystal faces, mica, selenite and quartz being examined. In this method a high-reflecting thin film of silver is deposited upon the crystal face which is then brought close to a similarly coated good optical flat. When illuminated under specified critical conditions with monochromatic light, sharpened multiple beam Fizeau 'fringes of equal thickness' are formed. These reveal surface topographical information, demonstrating on mica surfaces, for example, the existence of numerous small hills and dales and also of sharply defined cleavage steps often of molecular dimensions. In the case of mica and selenite, experimental conditions were such that there remained ambiguity as to whether a given contour group represented a hill or valley. Nor could the directions of the cleavage steps be decided. The contour picture given may equally well have been a mirror image of actual conditions. The fringes of equal chromatic order remove ambiguity and reveal further information of interest.

These fringes are essentially multiple-beam *white light* fringes, the basis of the precision attained lying in the employment of high-reflecting coefficients. The technique for the production of the necessary metallic films has already been described elsewhere. *An unusual feature is that the precision is reached without the* employment of a monochromatic light source, a 100 W lamp sufficing.

FRINGES OF EQUAL CHROMATIC ORDER

The formation of the fringes and their application may be examined by considering first those formed with a silver modified Newton's rings apparatus. It has been shown elsewhere (Tolansky 1943*b*) that the deposition of high-reflecting thin silvering upon the optical components (lens and glass plate) of a Newton's rings apparatus leads to a striking sharpening of the fringes when collimation is correct. Such sharpened fringes in transmission are shown in plate 12*a*. Local glass defects are rendered conspicuous and deviations of the order of 25 A can be measured with precision.

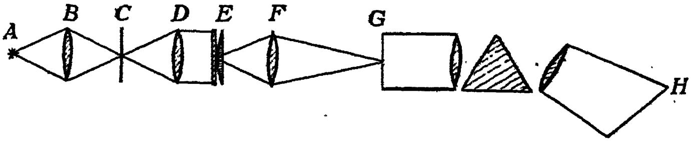


FIGURE 1

The arrangement for producing the fringes of equal chromatic order with this interference film is shown in figure 1. *A* is a white light source (a lamp), an image of which is projected by the lens *B* on the 1 mm. aperture *C*, which is at the focus of the achromat lens *D*. An approximately parallel beam falls at normal incidence upon the interference film *E*. An image of *E* (enlarged some 3–5 times) is projected by the good achromat lens F on to the narrow slit of a spectrograph G. The fringes are photographed at H.

The fringes covering the spectrum from the red to the blue are shown in plate 12*b*. A wave-length scale is superposed.

The fringes show the following characteristics:

(1) Although white light is used the fringes are very sharp, the fringe width being a small fraction of the distance between orders.

(2) Each fringe is multicoloured, extending over a considerable spectral range, which depends upon the focal length of the lens F.

(3) The approximately parabolic shape can be predicted from simple theory.

(4) The local defects visible in the Fizeau fringes on plate 12*a* appear as irregularities in the fringes on plate 12*b*.

(5) Each fringe is identical with the others, due allowance being made for the variable dispersion of the prism spectrograph employed.

(6) The separation between orders is large in terms of angstroms.

These fringes differ from, but are related to, the well-known Edser-Butler fringes formed with white light when a plane parallel 'half-silvered' thin air film is placed before the slit of a spectroscope. They are related also to the lesser known 'hook' fringes of Roschdestwensky. These relationships are discussed elsewhere. The following suffices to give the relationships required for the problems considered here.

The interference in the silver-enclosed film obeys the relation $n\lambda = 2\mu t \cos\theta$, where the symbols have the usual meanings. In this instance μ and $\cos\theta$ are taken to be unity, whilst t varies in a regular manner across the slit. This relation can be written as $n = 2\nu t$, where ν is the wave number. Differentiating for variation of n, with t constant, gives $dn = 2d\nu t$. For successive orders $dn = 1$, thus if $\Delta\nu$ is the wave-number separation between orders then, $t = 1/2\Delta\nu$. With a plane parallel film (t constant) the fringes are parallel lines equally spaced in wave numbers. Edser-Butler fringes correspond closely to this special case. When t varies across the slit image, then $\Delta\nu$ is greatest where t is smallest. If the point of closest approach in E is imaged on the slit centre, then it is clear, without detailed analysis, that parabolic shaped fringes will result.

For monochromatic fringes obeying the relation $n\lambda = 2\mu t \cos\theta$ Lummer long ago proposed the following classification. With a parallel-sided film (t constant) interference produces 'fringes of equal inclination'. With a wedge, using parallel light (θ constant) interference leads to 'fringes of equal thickness'. This terminology has received wide acceptance. In each of the fringes in plate 12*b* the condition obeyed is t/λ constant. This quantity, the order of interference (constant for any type of fringe), applies here to fringes which are *chromatic*, since λ varies. To correlate these with Lummer's classification, it is proposed to adopt the name '*fringes of equal chromatic order*'. A more detailed justification for this is being given elsewhere, where, in addition, the fringes formed by the simultaneous variations of both λ and θ are discussed.

Amongst the many possible uses of the fringes of equal chromatic order, the following applications will be surveyed here:

(1) Determination of hill and dale features.

(2) Measurement of small cleavage steps and determination of step direction.

(3) Measurement of very small surface angles.

(4) Evaluation of birefringence over a wide frequency range.

(5) Detection of small thickness differences, approaching molecular dimensions.

In each case a suitable modification of the formula $t = 1/2\Delta\nu$ enables precision measurements to be made rapidly. A fringe pattern as on plate 12*b* is effectively a section diagram of a surface contour, the section being that selected by the spectroscope slit. With a curved surface (the lens in *E*) definition is improved if the image projected on to the slit is enlarged. This is then equivalent to the examination of a thin slice of the surface narrower than that of the slit by the appropriate magnification factor. The improvement is due to the fact that the variation in film thickness across the *slit width* (despite the employment of a fine slit) is appreciable to the degree of sensitivity employed. There thus arises the unusual optical situation in which an enlargement in projection leads to improvement in definition.

It has been shown elsewhere that in monochromatic multiple-beam fringes the interference film must be very thin if high definition in the fringes is required. This condition also applies to the white light multiple-beam fringes. If the interference film is wedge-shaped, set with its edge perpendicular to the slit, the fringes are straight lines sloping across the slit image direction, the fringe slope being a measure of the wedge angle. Definition in this case is not affected by enlargement, since the film thickness varies only along the slit length and not across the slit width. Enlargement reduces the fringe slope, and it is helpful in the case of small wedge angles if the image is *reduced*, increasing thereby the fringe slope. With a plane parallel film, magnification or reduction of the image has no effect on definition. It is clear that a suitable choice of image magnification should be made, according to the problem. The guiding criterion is that magnification affects only the surface area, not the film thickness.

By sliding the image of the interference film across the slit, a large area can be scanned. This is rapidly done visually, but is photographically slow, the covering of an area being a more laborious process than is the case with Fizeau fringes.

A noteworthy feature is the attainment of high precision (approaching molecular dimensions) *without* the use of the monochromatic sources generally required for thin film interferometry.

In mica and selenite the fringes of equal chromatic order have been used in two types of experiment: (*a*) surface topography studies, (*b*) examination of internal structure features of thin slips. These studies resolve the ambiguities remaining in the earlier Fizeau fringe experiments.

APPLICATIONS TO SURFACE TOPOGRAPHY

The crystal surface (mica) is silvered and matched close to a silvered optical flat, as formerly. This combination replaces the lens-flat (E) in figure 1. The fringes given by a smooth sample of mica are shown in plate 12*c*, with an enlarged section in plate 12*d*. A surface cleavage splits the pattern in two. Comparison with the Newton's ring fringes shows by inspection that the upper half is a valley feature and the lower half a hillock.

Since there are more fringes above than below the dividing line, and as $t = 1/2\Delta\nu$, inspection suffices to show that the air film is thinner in the lower section, hence the direction of the cleavage step is determined. Precision values for cleavage steps and heights of hill features can be derived by a number of separate procedures. The following methods can be used for step evaluation.

For a particular order n, at some wave-length λ, the air film thickness being t, then $n\lambda = 2t$, giving $dt = d\lambda/2\lambda\Delta\nu$, where $\Delta\nu$ has the previous meaning and can be obtained from any fringe pair. It is usually sufficiently accurate to evaluate the wave-length displacement $d\lambda$ at the mean value λ to obtain dt. The method is suitable when the displacement is a small fraction of an order.

A better procedure is to measure the change in order at a given wave-length. In this case λ is constant and dn varies with t, thus $dt = dn\,\frac{1}{2}\lambda$, which is a convenient relationship to apply. The fractional order of separation dn is measurable with precision and inspection suffices to show whether an integer must be added, which is the case when orders have overlapped.

A further method is to make use of coincidences. Let m and m' be the number of fringes (including fractions) between any two selected points $\Delta\nu$ apart. If t and t' are the thicknesses of the two corresponding air films then $t = m/2\Delta\nu$ and $t' = m'/2\Delta\nu$, so that $t-t' = dt = (m-m')/2\Delta\nu$. In the case where $m = m'+1$ then $dt = 1/2\Delta\nu$, which occurs for perfect coincidences only.

Since the fringes are displaced above and below the dividing line, both resolution and precision of measurement exceed that normally possible with usual multiple-beam fringes of the same degree of sharpness, as, for example, Fabry-Perot rings. The condition here is analogous to that which obtains in the use of displaced fiduciary fringes such as are used in the Rayleigh refractometer. A displacement of 1/125th of an order can be accurately measured at 5000 A. This corresponds to a step of 20 A., *which is the molecular lattice spacing of mica*. The step shown is large, being some 16,000 A.

It may be noted that the smoothness of the fringes proves that mica cleaves true to a 'molecule'. This will be discussed elsewhere.

A further example of the use of the fringes of equal chromatic order is shown in plate 12*e*, an enlarged section is in plate 12*f*. This exhibits a sharp surface ridge, details of which are discussed elsewhere. It will be seen from this example that a useful characteristic of the fringes is that they reveal small surface angles. Consider a fringe inclined to the vertical, arising from two points on the mica surface L cm.

apart, the difference in air film thickness being dt. The small wedge angle on the surface (angle between normals) is then $s = dt/L$. Writing the fringe slope as $S = Kd\lambda/L$ gives $s = Svt/K$, in which K depends upon the dispersion of the spectroscope. For any fringe (n constant) $t = 1/2\Delta\nu$, so that $s = S\nu/2K\Delta\nu$. The magnification of the slope of the fringe over that of the surface wedge is thus $M = 2K\Delta\nu/\nu$. At 5000 A. a typically good value for $\Delta\nu$ is 1500 cm.$^{-1}$, and with the spectrograph in use the dispersion is some 30 A./mm., giving $M = 50{,}000$. With a Hilger large glass Littrow spectrograph this would be increased to over 500,000.

With the Hilger medium quartz spectrograph, used throughout here, a slope of $2\frac{1}{2}^\circ$ can be detected on a 1 cm. length of fringe. This corresponds to a wedge on the mica of $\frac{1}{20000}^\circ$, i.e. some 0·003 min. of arc. Comparison with the Rayleigh limit of the goniometer is of interest. At best, over a 1 cm. length, and providing too that a large enough area is available to give sufficient intensity in the reflected signal, the goniometer can resolve some $\frac{1}{400}^\circ$. The fringes of equal chromatic order are at least 50 times better than this, probably 100 times better. When small facets are involved the goniometer is far surpassed.

Further applications of the fringes will be indicated in the following section dealing with Fizeau fringes, where it will be shown how the two fringe types can be correlated.

FIZEAU FRINGES WITH DOUBLY SILVERED MICA

Since the topographical features of any single mica surface are always complex, it was anticipated that the multiple-beam monochromatic Fizeau fringes formed between the two faces of a single slip of mica would be more complex because of superposition. However, the effects observed are simpler than in the case of a single face matched against a flat. Experiments were made on thin rectangular slips with dimensions some 3 × 4 cm., thickness varying from $\frac{1}{150}$ to $\frac{1}{50}$ mm. Such thin flexible sheets were sensitive to air shock and in most of the observations were enclosed between glass plates held together with gummed paper strips. This leads to the formation of weak low-visibility secondary fringes formed in the air film between a glass surface and silvered mica surface. One air film only was involved, a large wedge being formed on the other face with a paper separator.

The doubly silvered mica was illuminated with parallel monochromatic light (green mercury) at normal incidence, using a small source at the principal focus of a good lens. To reduce complexity arising from the birefringence of mica, plane polarized light was used. The interference patterns which result are of interest and the visual appearance is striking if an unfiltered mercury arc is used as the source. Interference patterns (monochromatic) for different mica samples are shown in plate 13, each representing some 3 sq.cm. The sample shown in plate 13*j* exhibits moderately large areas of uniform tint separated by cleavage lines which radiate from the point of insertion of the needle used for separating the sheet. These cleavage lines can be seen on the mica surface with a low-power lens. The sample in plate 13*g* shows a mass of irregular torn cleavage strips of considerable complexity. Patterns between these two extremes have been observed in different samples.

Within the uniform tint areas are three major characteristic features:

(*a*) Gradual changes in depth of tint.

(*b*) Striations and irregular discontinuous patches.

(*c*) Secondary low-visibility fringes.

The characters (*a*) and (*b*) are due to changes in refractive index arising from variations in chemical composition. These aspects are being discussed elsewhere. The low-visibility secondary fringes still appear when the mica rests upon an optical flat and thus represent the contour of the surface topography of that mica surface in contact with the glass.

In plate 13*m* the *doubly* silvered mica rests upon a silvered optical flat. Compound Fizeau fringes result. The mica acts as a filter leading to uniform tint areas bounded by cleavage lines. Within each area of transmission sharp multiple-beam fringes form between the silvered glass and mica surfaces. Uniform tint shows that both faces of the mica are parallel to within a small fraction of a wave, yet the secondary fringes show that the surface is highly contorted. It must be concluded that the distortions on the one face are faithfully followed on the other face of the mica slip. Because of this the main features of the fringe pattern from a piece of doubly silvered mica are much less complex than those of either separate face when matched against a flat.

The 'parallel distortions' on the two faces do not prevent formation of uniform tint, which is determined by the condition $n\lambda = 2\mu t \cos\theta$. For t is postulated constant and the minute variations in θ due to the heights of the surface contours are negligible. Thus it is clear why Fabry-Perot fringes (or the original Haidinger rings) can be perfectly formed by mica with very imperfect surfaces.

CORRELATION WITH FRINGES OF EQUAL CHROMATIC ORDER

Owing to the local parallelism of the opposing sides of a slip of mica, the Fizeau fringes fail to give information about the directions and values of the steps, merely indicating their existence. The missing data are supplied by the fringes of equal chromatic order. These fringes corresponding to a line section 1·5 cm. in length crossing over one of the cleavage lines of plate 13*j* are shown alongside the latter in plate 13*k* (enlarged section in plate 13*i*). The fringes are straight lines, parallel to the slit, affording independent proof that the mica faces are parallel. The fringes are double, because of the birefringence of the mica.

These fringes reveal many properties, e.g.

(*a*) The direction of the cleavage step.

(*b*) The height of the step. It is 120 A, i.e. exactly 6 'molecules'.

(*c*) The thickness of the mica. It is $\frac{1}{59}$ mm.

(*d*) Secondary intensity variations superposed on the whole pattern.

(*e*) The birefringence over an extended wave-length range.

The secondary fringes reveal the surface topography, since they are the low-visibility fringes of equal chromatic order formed between the mica surface and the unsilvered flat on which this rests. This again proves that the surface is contorted, yet parallelism retained. There exists point to point correspondence between the two types of fringes.

The fringes of equal chromatic order shown in plate 13*h* (enlarged section in plate 13*i*) refer to a 1·4 cm. length of part of the sample shown in plate 13*g*. Again correspondence can be traced, but it is clear that the fringes of equal chromatic order reveal much additional detail about the cleavage steps. One such conclusion, for example, is that changes in thickness are always discontinuous (excluding some special local features associated with inclusions and growth nuclei).

The details of the analysis of the interference pictures are not being given here, since the object in view is the description of methods. It can, however, be demonstrated that steps occur which are all small integral values of 20 A, which is the crystal lattice spacing of mica in the direction concerned with cleavage.

A special characteristic of doubly silvered mica fringes is that the 'optical' thickness being involved, displacements are magnified by the factor μ. The sensitivity is thus some 60 % better than when using an air film as in the topographical studies on single surfaces matched against a flat.

Finally, the method has been applied to the study of small local features involving what are considered to be growth nuclei and occlusions. The results of these investigations are being communicated elsewhere.

BIREFRINGENCE OF CLEAVED MICA SHEETS

The fringes of equal chromatic order permit a determination of birefringence to be made over a wide wave-length range from a single exposure. The refractive indices of micas differ, determinations for any given sample not necessarily being representative. The average three refractive indices (*D* lines) given by Rutley for muscovite (*Mineralogy*, 23rd ed.) are 1·560, 1·593, 1·600. The *International Critical Tables* give 1·561, 1·590, 1·594 for muscovite. The birefringence expected in a cleavage flake is thus small, of the order 0·007–0·004.

Consider the wave-length change $d\lambda$ for a fringe of given order, due to a change in refractive index $d\mu$. As $n\lambda = 2\mu t$ differentiation gives $nd\lambda = 2d\mu t$, hence the birefringence $d\mu$ is given by $d\mu = \mu d\lambda/\lambda$. This can be measured over a wide spectral range by evaluating the doublet separations $d\lambda$, the value of μ for the wave-lengths selected being obtained from the wave-number separation between any two adjacent orders,* or by other means.

Measurement shows that $d\mu$ is sensibly constant over the visible region, having the value 0·0046 ± 0·0001 for the sample examined, at 5000 A.

---

* Strictly, taking into account dispersion, then $1/2\mu t = \Delta\nu + \nu\Delta\mu/\mu$, where $\Delta\mu$ is the change in μ over the range $\Delta\nu$. For an approximate evaluation of μ, $1/2\mu t = \Delta\nu$ can be used.

Since $d\lambda = d\mu\lambda/\mu$, the birefringent doubling, in Angstrom units, is independent of the thickness of the mica film. This has been confirmed by observation. It is estimated that a birefringence of 0·0001 could be measured by first polarizing the incident light and then separately independently measuring the positions of the doublet components. Even this low value could be improved upon in special instances with a suitable spectrograph.

NON-LOCALIZED FRINGES OF EQUAL INCLINATION

The third interferometric method developed for the examination of doubly silvered mica employs the high dispersion non-localized multiple-beam circular monochromatic fringes described recently by the writer (Tolansky 1943 *b*).

It is known that with selected pieces of doubly silvered mica, good Fabry-Perot rings can be formed if a small area is illuminated. Haidinger's original discovery of the fringes of equal inclination (1854) was made with unsilvered mica. Rayleigh (1906) was the first to silver (crudely) the two faces of a mica slip and he succeeded in resolving two ring systems, due to birefringence. Chinmayanandam (1919) and Schäfer & Herber (1926) extended these observations, and more recently Pfund (1942) has added more, without, it seems, being aware of the earlier literature.

The non-localized fringes formerly described by the writer are closely allied to true Fabry-Perot rings, although the latter are at infinity. The theory of their formation has already been discussed. It was shown that with a small source a Fabry-Perot interferometer produces characteristically highly sharpened rings which extend indefinitely into space, growing larger as the distance between the screen and source is increased. No lenses are required for the production of these fringes, which, providing certain conditions are fulfilled, are as sharp as Fabry-Perot rings. The most striking characteristic of these fringes is the high dispersion available with such simple means. Rings many metres in diameter can be obtained. If thin slips of mica are used as the interference film, it is necessary to tilt off-centre to accommodate more than one order upon a quarter plate placed only 10 cm. distant from the source.

The optical arrangement is simple. A small aperture is placed before a green-filtered mercury arc, and with a 1 in. microscope objective a diminished bright image of this circular aperture is focused on to the mica. The image size is some 0·25 mm. across. The mica is tilted to bring down higher orders and a photographic plate held some 20 cm. distant. Two successive orders of the rings formed when an image of the source is thrown on to the uniform tint area of plate 13 *j* are shown in plate 14 *o*. The light is first suitably plane polarized to cut out one ring system.

The rings are sharp and the high dispersion is evident. The linear distance on the plate between orders is 3 cm. It would have been made 30 cm. by moving the plate 2 m. from the mica instead of 20 cm. With green mercury light the change in mica thickness needed to displace the rings through one order is 1700 A. It is evident that a change of less than 1/100th order can be measured, i.e. a step of 17 A or less can be detected, whilst a setting on a fringe can be made to within 1/1000th order. The sensitivity is thus the same as with the fringes of equal chromatic order.

For comparison the true Fabry-Perot fringes (at infinity) are shown in plate 14*n*. This mica thickness is so uniform over a large region that a considerable surface area can be illuminated with little loss in definition. A 6 mm. disk was here illuminated. The fringe sharpness again confirms the earlier conclusion about parallelism of the sides despite surface contortion.

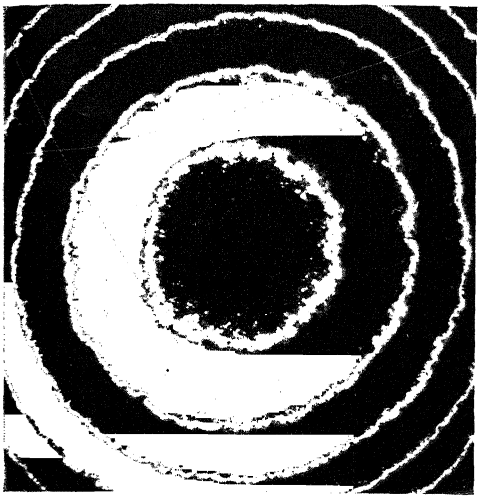
The non-localized fringes (with a small source image) shown in plate 14*p* were obtained when the source image was placed to cross over one of the cleavage lines in plate 13*j*. A discontinuity appears in the fringe pattern. There exist effectively two Fabry-Perot interferometers, side by side, differing in gap only by the height of the cleavage step. The incident converging beam is divided by the cleavage line so that on one side of the line some half of the converging cone is available and therefore only half the ring system appears. On the other side of the cleavage line is the other half of the ring system, from the second interferometer, with a displacement in order corresponding to the difference in gap. From the displacement a precision value of the step can be obtained.

This one photograph is insufficient to give the step, since there is no indication either of the direction of the displacement or of the number of integers to be added to the observed fraction. There are two methods for resolving the ambiguity. The first is to use different wave-lengths and then apply Benoit's method of exact fractions. This is laborious and requires a knowledge of the dispersion. The second is simpler. The fringes are compared with the corresponding fringes of equal chromatic order, and this reveals whether an integer is to be added. The one advantage of these non-localized fringes over those of equal chromatic order is the great dispersion of the former, obtained with simple means.

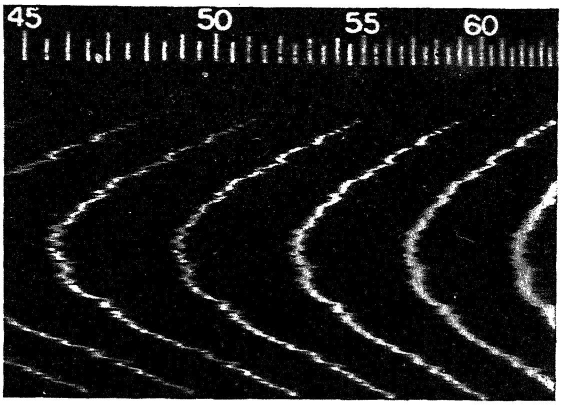
Plate 14*q* was taken over the same region as plate 14*p* but with the polarizing device removed. The birefringence doubling appears, one ring system being circular, the other slightly elliptical. This effect was indicated, although much less clearly, in the earlier work of Rayleigh on the Fabry-Perot rings and was also clearly established by Pfund with Fabry-Perot rings. None of the earlier workers who studied the Fabry-Perot rings succeeded in discovering the step effects rendered so easily visible with the non-localized fringes.

The non-localized fringes on plates 14*r* and 14*s* were obtained from small areas of the mica sample in plate 13*g*, which has a torn surface. In plate 14*r* the small source image is covering an area containing a number of discrete patches extending horizontally over sufficient length to encompass a considerable angular range for each strip. In plate 14*s* the strips lie more or less vertically, leading to isolated fragments of fringes. In both cases the light was first plane polarized to reduce complexity.

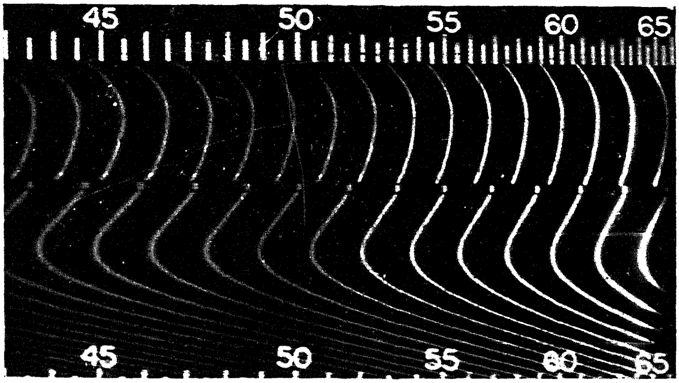
(*Historical note.*) On completion of this work a search was made through the extensive literature on multiple-beam interferometry to find whether the fringes (called here fringes of equal chromatic order) have been discussed before. Nothing has been found. The closest approach appears to be in the work of Fabry & Buisson



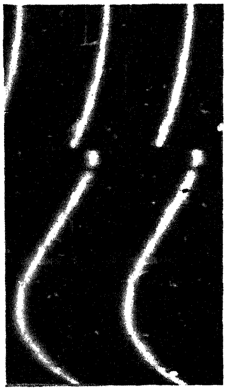
a



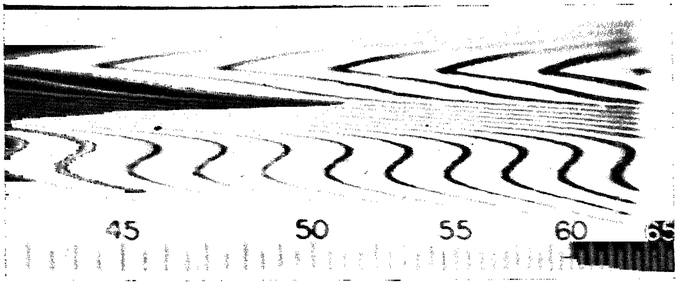
b



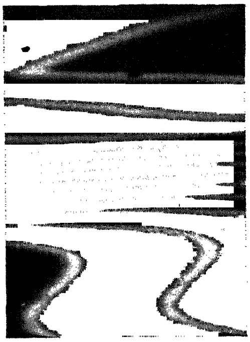
c



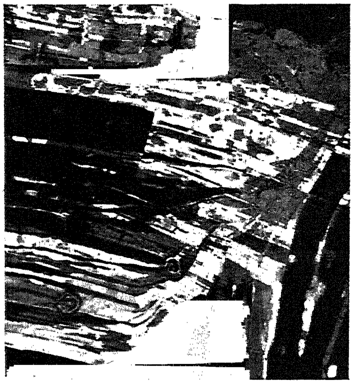
d



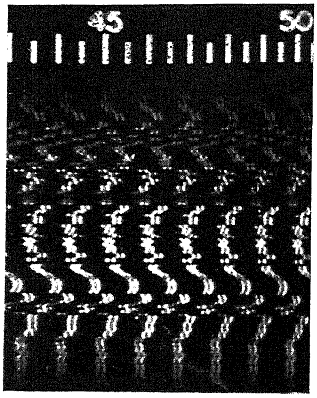
e



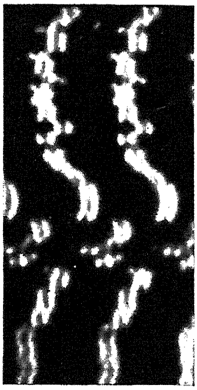
f



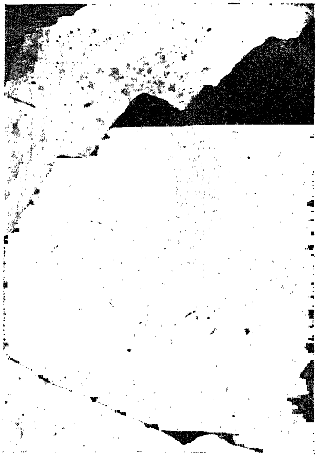
g



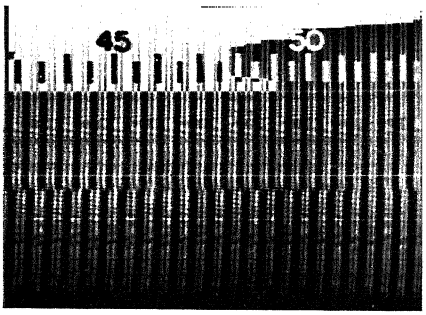
h



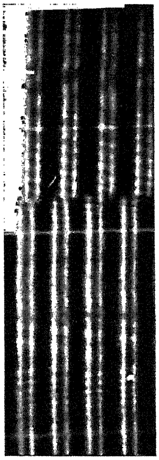
i



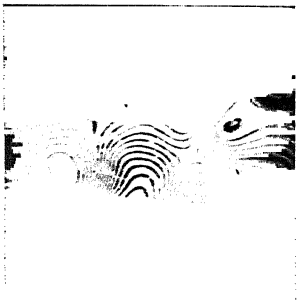
j



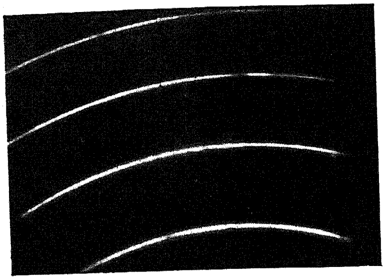
k



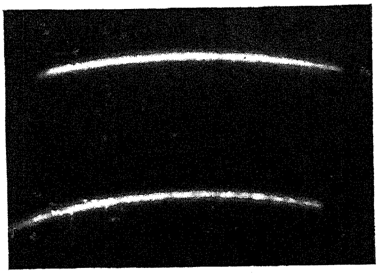
l



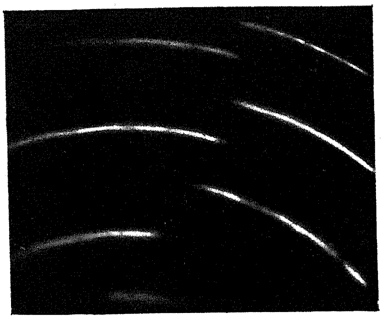
m



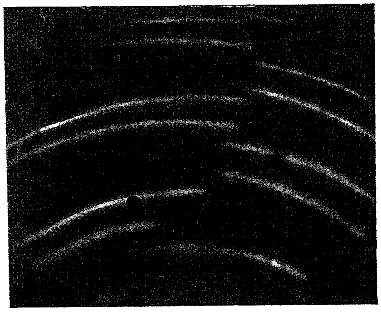
n



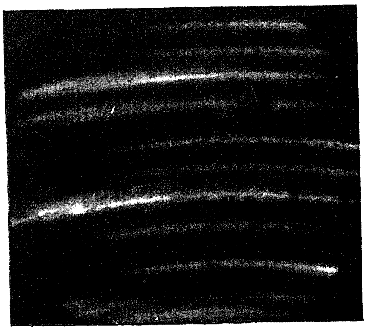
o



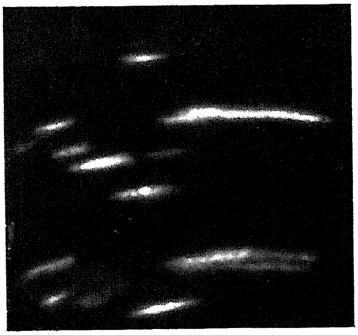
p



q



r



s

(1908) in their determination of phase change on reflexion. They examined the 'channel' spectrum given by a wedge, but it appears from their brief description that the optical conditions necessary for the formation of the fringes of equal chromatic order were *not* realized, and they abandoned their observations because of the mediocre fringes formed when white light was used. Without stating so, they seem to imply that the fringes examined were of the Edser-Butler type. The fringe types recounted here were not described.

REFERENCES

Chinmayanandam 1919 *Proc. Roy. Soc.* A, **95**, 176.
Fabry & Buisson 1908 *J. Phys. Radium*, **4**, 417.
Pfund 1942 *J. Opt. Soc. Amer.* **32**, 383.
Rayleigh 1906 *Phil. Mag.* **12**, 489.
Schäfer & Herber 1926 *Z. tech. Phys.* **7**, 98.
Tolansky 1943*a* *Nature*, **152**, 722.
Tolansky 1943*b* *Phil. Mag.* **34**, 555.
Tolansky 1944*a* *Nature*, **153**, 195.
Tolansky 1944*b* *Phil. Mag.* **35**, 120.
Tolansky 1945 *Proc. Roy. Soc.* A, **184**, 41.

The air wave surrounding an expanding sphere

By G. I. Taylor, F.R.S.

(*Received* 5 *December* 1939)

[For summary see p. 292.]

Introduction

When the surface of a sphere vibrates in any assigned manner the spherical sound waves which are propagated outwards can be represented by well-known formulae provided that the motion is such that only small changes in air density occur. When the motion of the spherical surface is radial the velocity potential of the sound wave is

$$\phi = r^{-1}f(r-at), \tag{1}$$

where a is the velocity of sound and r is the radial co-ordinate. The velocity, u, and the excess, $p-p_0$, of pressure over the atmospheric pressure p_0 are

$$u = r^{-2}f(r-at) - r^{-1}f'(r-at), \tag{2}$$

$$p-p_0 = -\rho ar^{-1}f'(r-at). \tag{3}$$

If R is the radius of the sphere which, by its expansion, is producing waves, R is a function of t and the surface condition is

$$\dot{R} = R^{-2}f(R-at) - R^{-1}f'(R-at). \tag{4}$$

Equation (4) is an equation for finding the function f. A simple case in which equation (4) can be solved is when $\dot{R}$ is constant so that the sphere is expanding at a uniform velocity. Taking $t=0$ when $R=0$ the radius at time t can be expressed in the form

$$R = \alpha at, \tag{5}$$

where α is a non-dimensional constant. The limitation that the changes in density are small implies that equations (1)–(3) are true only where α is small compared with 1.

Writing $w = R-at = (\alpha-1)\,at$ equation (5) becomes

$$\frac{\alpha-1}{\alpha w}f'(w) - \left(\frac{\alpha-1}{\alpha w}\right)^2 f(w) + a\alpha = 0. \tag{6}$$

The solution of equation (6) which is valid for negative values of w is

$$f(w) = \frac{a\alpha^3}{1-\alpha^2}w^2 + c(-w)^{(\alpha-1)/\alpha}. \tag{7}$$

The constant of integration c must be taken as zero in order that $f(w)$ may vanish when $w = 0$. Hence

$$\phi = \frac{a\alpha^3}{1-\alpha^2}\frac{(r-at)^2}{r}, \tag{8}$$

$$\left.\begin{aligned} u &= \frac{a\alpha^3}{1-\alpha^2}\left(\frac{a^2t^2}{r^2}-1\right), \\ p-p_0 &= 2\rho\frac{a^2\alpha^3}{1-\alpha^2}\left(\frac{at}{r}-1\right). \end{aligned}\right\} \tag{9}$$

If at time $t = 0$, $u = 0$ and $p-p_0 = 0$ everywhere, then at all subsequent times $u = 0$ and $p = p_0$ in the region outside the sphere $r = at$.

It will be seen that both u and $p-p_0$ are constant when r/at is constant, thus points where u and $p-p_0$ have any assigned value are propagated outwards at uniform speeds which are proportional to distance from the centre. Subject to the limitations of the theory of sound therefore* the air wave produced by a uniformly expanding sphere expands at a uniform rate and the velocity and pressure at corresponding points are constant at all stages of the expansion.

This result might have been expected *a priori* but the solution is here given in detail because it forms the starting point of the work which follows.

ANALYSIS WHEN VELOCITY OF EXPANSION IS NOT SMALL

It seems likely that a uniformly expanding sphere will be surrounded by a uniformly expanding air wave, accordingly a solution of the complete equations of motion is sought in which u and p are functions of $x = r/t$ only. For such motions

$$\left(\frac{\partial}{\partial t}+\frac{r}{t}\frac{\partial}{\partial r}\right)(u, p, \rho) = 0. \tag{10}$$

The equation of motion is

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial r} = -\frac{1}{\rho}\frac{\partial p}{\partial r}, \tag{11}$$

and in view of equation (10) this may be written

$$(u-x)\frac{du}{dx} = -\frac{1}{\rho}\frac{dp}{dx}. \tag{12}$$

---

* It is shown later that the sound wave equations themselves are not valid in this case, even when α is small, but this fact does not invalidate the expression (8) regarded as a solution of those equations.

The equation of continuity is

$$\frac{\partial\rho}{\partial t}+u\frac{\partial\rho}{\partial r}+\rho\left(\frac{\partial u}{\partial r}+\frac{2u}{r}\right)=0,$$

which in view of equation (10) may be written

$$\frac{u-x}{\rho}\frac{d\rho}{dx}+\frac{du}{dx}+\frac{2u}{x}. \tag{13}$$

The gas equation $p\rho^{-\gamma}$ = constant, together with the expression for the velocity of sound, namely $c^2=\frac{dp}{d\rho}=\frac{\gamma p}{\rho}$, give

$$\left.\begin{aligned}\frac{1}{\rho}\frac{dp}{dx}&=\frac{1}{\gamma-1}\frac{dc^2}{dx},\\ \text{and}\qquad \frac{1}{\rho}\frac{d\rho}{dx}&=\frac{1}{(\gamma-1)c^2}\frac{dc^2}{dx}.\end{aligned}\right\} \tag{14}$$

Substituting from equations (14) in equations (12) and (13)

$$\frac{dc^2}{dx}=-(\gamma-1)(u-x)\frac{du}{dx}, \tag{15}$$

$$\frac{u-x}{(\gamma-1)c^2}\frac{dc^2}{dx}+\frac{du}{dx}+\frac{2u}{x}=0. \tag{16}$$

For convenience in calculation equation (16) may be replaced by

$$\frac{du}{dx}=-\frac{2u}{x}\left\{1-\left(\frac{u-x}{c}\right)^2\right\}^{-1}. \tag{17}$$

Equations (15) and (17) may be expressed in non-dimensional form by substituting the variables

$$\left.\begin{aligned}\xi&=u/x,\\ \eta&=c^2/x^2,\\ z&=\log_e x.\end{aligned}\right\} \tag{18}$$

The resulting equations are

$$\frac{d\eta}{d\xi}=\frac{2\eta}{\xi}\frac{\eta+(\gamma+1)\xi-\gamma\xi^2-1}{3\eta-(1-\xi)^2}, \tag{19}$$

$$\frac{dz}{d\xi}=-\frac{1}{\xi}\frac{\eta-(1-\xi)^2}{3\eta-(1-\xi)^2}. \tag{20}$$

The solution of equation (19), which contains two variables only, will contain one arbitrary constant. Without attempting to express this solution in mathematical form it is possible to construct by numerical integration a complete set of (ξ, η) relationships each corresponding with a given value of the arbitrary constant and to set them out graphically in a single set of curves on a diagram whose co-ordinates are ξ and η. This diagram is shown in figure 1. The arbitrary constant α is defined so as to correspond with the constant α in equations (5), (8) and (9), and the value of α corresponding with each (ξ, η) curve is shown in figure 1. The single curve which cuts across all the graphical solutions of equation (19) in figure 1 will be explained later.

BOUNDARY CONDITION AT THE SURFACE OF THE SPHERE

At the surface of the expanding sphere $u = r/t = x$ so that

$$\xi = 1. \tag{21}$$

In the sound-wave solution the constant α specifies the velocity of radial expansion of the sphere as a fraction of the velocity of sound. In the complete solution η represents c^2/x^2 at any point so that at the surface of the sphere η represents $\left(\frac{\text{local velocity of sound}}{\text{velocity of expansion}}\right)^2$. Correspondence between the complete solution and the sound-wave solution is therefore attained when the arbitrary constant α is defined by the relation

$$\eta_2 = \alpha^{-2}, \tag{22}$$

where η_2 is the value of η at $\xi = 1$.

The curves in figure 1 were constructed for a series of values of α starting at the point $(\xi = 1, \eta = \alpha^{-2})$ and calculating the change in η step by step for small decrements $\delta\xi$ in ξ through the range $\xi = 1$ to $\xi = 0$. The change δz in z in each interval $\delta\xi$ was also calculated using equation (20). When η has been found as a function of ξ the solution of equation (20) is of the form

$$z - z_2 = (\text{function of } \xi),$$

where z_2 is the constant of integration. If z_2 is chosen so that $z = z_2$ when $\xi = 1$ then

$$\Sigma \delta z = z - z_2 = \log_e r/R. \tag{23}$$

OUTER BOUNDARY CONDITIONS

At the outer boundary of the expanding air it must be possible to connect the still air conditions with those which obtain in the disturbed region. This can be done in two possible ways: either (*a*) it might be found that $u = 0$ at radius $r = at$, so that the (ξ, η) curve passes through the point $(\xi = 0, \eta = 1)$; or (*b*) it might be found that at some radius the pressure temperature and velocity are attained which correspond with the pressure, temperature and velocity immediately behind a shock wave moving into still air with velocity r/t. In case (b) the expanding region would be bounded by an expanding spherical shock wave and the air outside this sphere would be at rest. The sound wave solution (equations (8) and (9)) satisfies condition (a), for at $r = at$ both $u = 0$ and $p - p_0 = 0$.

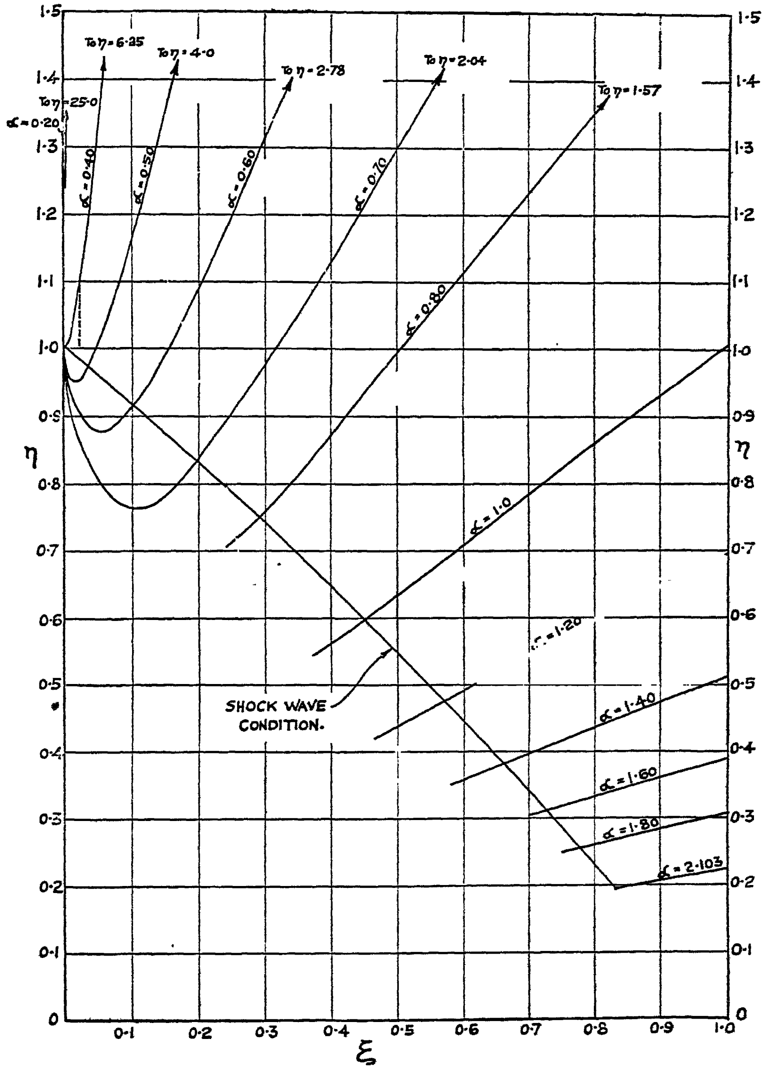


FIGURE 1*

---

* The area of figure 9 is indicated by a broken line on the right of the area.

For values of α larger than those to which sound wave analysis can be applied it appears that it is not possible to satisfy condition (*a*). If the solution for $\alpha = 0{\cdot}7$ for instance is followed (see figure 1) for values of ξ decreasing from $\xi = 1$ it is found that z reaches a maximum while ξ is still positive and for smaller values of ξ, z decreases. Thus the same value of z would correspond with two different values of ξ which is physically impossible. It remains to find out whether the alternative condition (*b*) can be satisfied. For this purpose it is necessary to express the appropriate shock wave conditions in terms of the variables ξ and η.

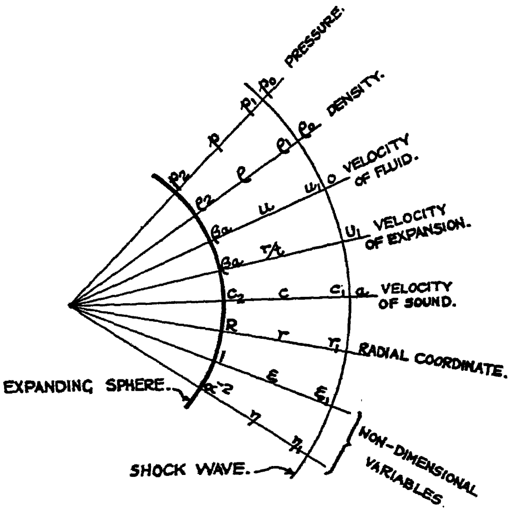


FIGURE 2. Diagram of symbols.

A shock wave is an extremely thin region within which the pressure, density, temperature and velocity change from one set of values to another. The ratios of density and temperature on the two sides of a shock wave depend only on the ratio of the corresponding pressures. If p_1 is the pressure immediately behind a shock wave and p_0 is the atmospheric pressure in front of it, $y = p_1/p_0$ may be regarded as the independent variable in terms of which all other changes occurring at the shock wave may be expressed. Figure 2 shows the positions in the field to which the various symbols apply. The shock wave formulae were first given by Rankine (1870) and later independently by Hugoniot (1889). The ratio of densities on the two sides of the shock wave is

$$\frac{\rho_1}{\rho_0} = \frac{\gamma - 1 + (\gamma + 1)\,y}{\gamma + 1 + (\gamma - 1)\,y}, \tag{24}$$

where γ, the ratio of specific heats, is the same as the "γ" which appears in equation (19). In the present calculations γ is taken as 1·405. Continuity requires that ρ_1/ρ_0 must be equal to the ratio of the velocities of the air on the two sides relative to the shock wave itself. Hence if U_1 is the velocity of the shock wave and u_1 that of the air behind it

$$\frac{U_1-u_1}{U_1}=\frac{\rho_0}{\rho_1},$$

or, using equation (24), $$\frac{u_1}{U_1}=\frac{2(y-1)}{\gamma-1+(\gamma+1)y}. \tag{25}$$

The condition that the shock wave may expand uniformly with the rest of the system is

$$U_1=r_1/t,$$

where r_1 is the radius of the shock wave.

Hence $$u_1/U_1=u_1t/r=\xi_1,$$

or $$\xi_1=\frac{2(y-1)}{\gamma-1+(\gamma+1)y}. \tag{26}$$

So far as this condition is concerned an appropriate value of y may be chosen and a corresponding possible shock wave found at any point in the field. Equation (26), however, is not the only necessary condition. The velocity of sound in the air behind the shock wave must also satisfy the condition

$$\eta_1=c_1^2t^2/r_1^2=c_1^2/U_1^2, \tag{27}$$

and $$c_1^2=\gamma p_1/\rho_1=a^2y\left\{\frac{\gamma+1+(\gamma-1)y}{\gamma-1+(\gamma+1)y}\right\}, \tag{28}$$

where a is the velocity of sound in the undisturbed atmosphere. U_1 may be expressed in terms of a making use of the momentum equation

$$\rho_0U_1u_1=p_1-p_0=(y-1)p_0, \tag{29}$$

substituting a^2/γ for p_0/ρ_0 and u_1/U_1 from (25)

$$\frac{a^2}{U_1^2}=\frac{2\gamma}{\gamma-1+(\gamma+1)y}, \tag{30}$$

hence from equations (27), (28) and (30)

$$\eta_1=\frac{2\gamma y\{\gamma+1+(\gamma-1)y\}}{\{\gamma-1+(\gamma+1)y\}^2}. \tag{31}$$

Values of ξ and η have been calculated from equations (26) and (31) for a sequence of values of y. These are plotted in a curve in figure 1. The intersections of this curve with those which describe the flow in the expanding air determine the values of ξ and η, and hence the value of y, corresponding with possible shock waves. These have been taken from figure 1 and are given in cols. 2 and 3 of table 1 for values of α ranging from 0·5 to 2·1. The corresponding values of r_1/R obtained by numerical integration of equation (20) are also given in col. 4 of table 1. Corresponding values of y are given in col. 5 of table 1.

TABLE 1

1 α	2 ξ_1	3 η_1	4 r_1/R	5 y	6 β	7 y'	8 p_2/p_0
0	—	—	—	—	—	—	—
0·2	—	—	4·93	1·000	0·203	0·928	1·075
0·4	0·0021	0·998	2·44	1·003	0·410	0·775	1·295
0·5	0·033	0·974	1·950	1·050	0·523	0·750	1·400
0·6	0·103	0·916	1·763	1·169	0·638	0·749	1·569
0·7	0·198	0·833	1·503	1·365	0·761	0·755	1·808
0·8	0·291	0·749	1·392	1·629	0·891	0·774	2·105
1·0	0·453	0·597	1·256	2·400	1·180	0·811	2·959
1·2	0·575	0·474	1·182	3·59	1·520	0·847	4·250
1·4	0·662	0·382	1·135	5·60	1·953	0·887	6·32
1·6	0·727	0·313	1·103	9·06	2·560	0·917	9·89
1·8	0·779	0·256	1·083	17·95	3·598	0·92	19·7
2·1	0·832	0·197	1·060	∞	∞	0·93	∞

REDUCTION TO MORE FAMILIAR FORMS

The physical cause of the motion of the expanding air being the motion of the sphere, the results are more comprehensible when expressed in terms of the ratio

$$\beta = \frac{\text{velocity of expanding spherical surface}}{\text{velocity of sound in undisturbed atmosphere}} = \frac{U_2}{a}, \tag{32}$$

rather than in terms of α. Since $U_2 t/U_1 t$ is R/r_1

$$\beta = \frac{U_2}{U_1}\frac{U_1}{a} = \frac{R}{r_1}\sqrt{\left(\frac{\gamma-1+(\gamma+1)y}{2\gamma}\right)}, \tag{33}$$

values of β are given in col. 6 of table 1.

The ratio $$y' = \frac{\text{pressure behind shock wave}}{\text{pressure at surface of sphere}} = \frac{p_1}{p_2}, \tag{34}$$

is related to c_1/c_2 by the equation

$$y'^{(1-\gamma)/\gamma} = c_2^2/c_1^2. \tag{35}$$

Values of y' are given in col. 7 of table 1. By definition of η

$$\frac{c_2^2}{c_1^2} = \frac{\eta_2 R^2}{\eta_1 r_1^2} = \frac{R^2}{r_1^2 \alpha^2 \eta_1},$$

so that
$$\frac{p_2}{p_0} = \frac{y}{y'} \doteq y\left(\frac{R^2}{r_1^2 \alpha^2 \eta_1}\right)^{\gamma/(\gamma-1)}. \tag{36}$$

Values of p_2/p_0 namely the pressure at the surface of the sphere expressed in atmospheres, are given in col. 8 of table 1. The way in which p_1/p_0 and p_2/p_0 vary with β is shown in figure 3.

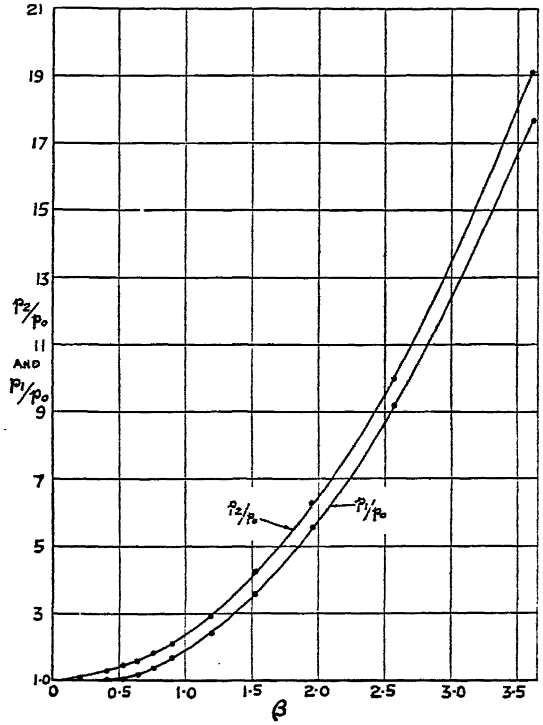


FIGURE 3. Pressure p_2 at sphere and p_1 behind shock wave as multiples of atmospheric pressure p_0.

LIMITING VALUES FOR VERY HIGH RATES OF EXPANSION

The limiting values of ξ_1 and η_1 when $y \to \infty$ are

$$\left.\begin{aligned} \xi &= \frac{2}{\gamma+1} = 0{\cdot}8316, \\ \eta &= \frac{2\gamma(\gamma-1)}{(\gamma+1)^2}. \end{aligned}\right\} \tag{37}$$

Starting from these values equations (19) and (20) were integrated numerically for increasing values of ξ. At $\xi = 1$ the value of η so found was 0·226 corresponding with $\alpha = (0{\cdot}226)^{-2} = 2{\cdot}103$, and the value of r_1/R was 1·0602. These are given in the last line of table 1. It seems therefore that as the velocity of expansion becomes infinitely great the pressure at the surface becomes infinite, but the density remains finite. The thickness of the layer of expanding air is never less than 6·0 % of the radius of the sphere.

VARIATION OF VELOCITY WITH RADIUS

The numerical solutions of equations (19) and (20) give η and r/R in terms of ξ. The most convenient variables for describing velocity distribution are u/a and r/at which are connected with ξ, η and r/R by the relations

$$\left.\begin{aligned} u/a &= \beta\xi r/R, \\ r/at &= \beta r/R. \end{aligned}\right\} \tag{38}$$

Some calculated values of u/a and r/at are given in table 2. These velocity distributions are shown in figure 4 for values of β ranging from 0·20 to 1·95.

TABLE 2

$\alpha = 0{\cdot}2$, $\beta = 0{\cdot}203$

r/at	u/a	p/p_0
0·203	0·203	1·0752
0·214	0·182	1·0745
0·228	0·159	1·0727
0·253	0·127	1·0671
0·300	0·090	1·0571
0·374	0·356	1·0431
0·425	0·042	1·0362
0·594	0·018	1·0196
0·766	0·008	1·0087
1·000	0·000	1·0000

$\alpha = 0{\cdot}4$, $\beta = 0{\cdot}410$

r/at	u/a	p/p_0
0·410	0·410	1·295
0·430	0·369	1·293
0·451	0·334	1·286
0·471	0·303	1·280
0·512	0·211	1·263
0·614	0·162	1·213
0·697	0·113	1·173
0·799	0·069	1·122
0·901	0·035	1·068
0·984		1·015
1·000		1·003

$\alpha = 0{\cdot}5$, $\beta = 0{\cdot}523$

r/at	u/a	p/p_0
0·523	0·523	1·400
0·544	0·481	1·397
0·564	0·444	1·391
0·586	0·411	1·386
0·627	0·353	1·363
0·669	0·304	1·338
0·711	0·262	1·310
0·774	0·209	1·265
0·836	0·162	1·219
0·900	0·120	1·171
0·940	0·093	1·137
0·983	0·065	1·100
1·017	0·031	1·050

$\alpha = 0{\cdot}6$, $\beta = 0{\cdot}638$

r/at	u/a	p/p_0
0·638	0·638	1·569
0·660	0·597	1·566
0·723	0·489	1·539
0·787	0·405	1·494
0·850	0·332	1·437
0·936	0·249	1·349
0·978	0·209	1·300
1·020	0·167	1·245
1·042	0·145	1·186
1·067	0·114	1·169

$\alpha = 0{\cdot}7$, $\beta = 0{\cdot}761$

r/at	u/a	p/p_0
0·761	0·761	1·808
0·782	0·717	
0·826	0·640	1·786
0·890	0·541	1·736
0·934	0·484	1·692
1·000	0·404	1·612
1·043	0·353	1·550
1·087	0·302	1·480
1·109	0·275	1·442
1·130	0·240	1·399
1·145	0·225	1·365

$\alpha = 0{\cdot}8$, $\beta = 0{\cdot}891$

r/at	u/a	p/p_0
0·891	0·891	2·105
0·935	0·805	2·096
0·980	0·729	2·067
1·025	0·662	2·022
1·068	0·598	1·965
1·114	0·537	1·898
1·158	0·478	1·820
1·203	0·417	1·733
1·225	0·384	1·677
1·242	0·357	1·630

TABLE 2 (*continued*)

$\alpha = 1{\cdot}0,\ \beta = 1{\cdot}180$

r/at	u/a	p/p_0
1·180	1·180	2·959
1·227	1·088	2·939
1·274	1·004	2·889
1·321	0·927	2·822
1·368	0·853	2·731
1·415	0·779	2·621
1·463	0·704	2·485
1·482	0·670	2·413

$\alpha = 1{\cdot}2,\ \beta = 1{\cdot}520$

r/at	u/a	p/p_0
1·520	1·520	4·250
1·570	1·421	4·231
1·620	1·330	4·169
1·671	1·233	4·067
1·722	1·157	3·927
1·772	1·071	3·747
1·800	1·029	3·636

$\alpha = 1{\cdot}4,\ \beta = 1{\cdot}953$

r/at	u/a	p/p_0
1·953	1·953	6·317
2·010	1·843	6·286
2·065	1·742	6·191
2·120	1·643	6·033
2·175	1·544	5·811
2·215	1·470	5·607

$\alpha = 1{\cdot}6,\ \beta = 2{\cdot}560$

r/at	u/a	p/p_0
2·560	2·560	9·89
2·603	2·474	9·87
2·649	2·385	9·82
2·696	2·291	9·72
2·750	2·198	9·50
2·824	2·050	9·07

$\alpha = 1{\cdot}8,\ \beta = 3{\cdot}60$

r/at	u/a	p/p_0
3·60	3·60	19·7
3·66	3·47	19·7
3·73	3·35	19·5
3·79	3·22	19·0
3·86	3·09	18·3
3·90	3·03	17·9

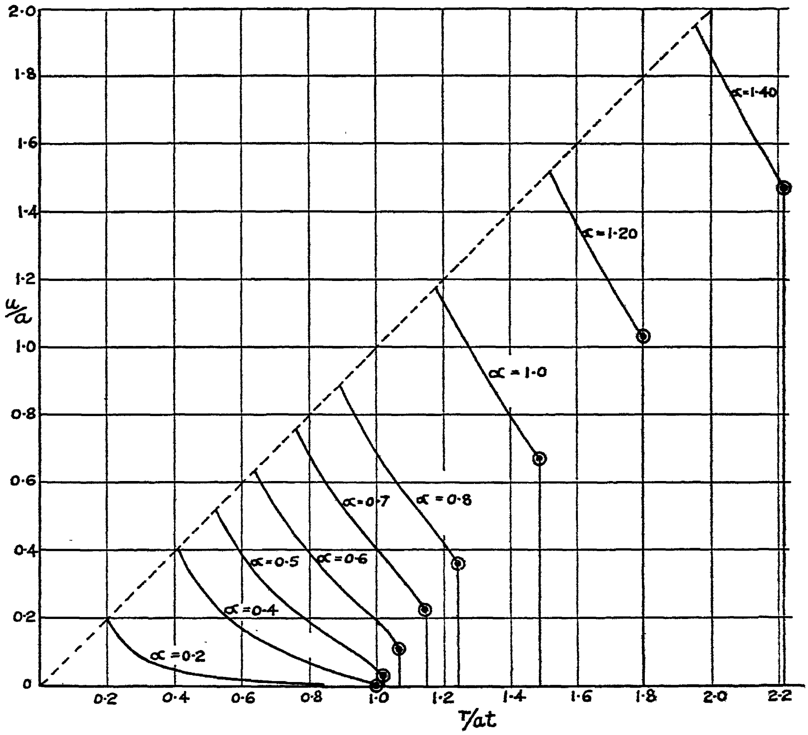


FIGURE 4. Distribution of velocity.

The sudden jump in velocity which occurs at the shock wave is represented in each case by a vertical line and the subsequent increase from the shock wave to the sphere by a sloping curve behind it. The points corresponding with the surface of the expanding sphere lie on the line $u = r/t$ because this is the condition which must be satisfied at the sphere. For high rates of expansion the velocity distribution is practically linear. When the thickness of the layer of expanding air is small compared with the radius of the sphere $\frac{1}{t}\frac{du}{dr}$ does not differ appreciably from its value at the spherical surface which is, according to equation (17), equal to -2. The mean slope of the velocity distribution curve for $\beta = 1{\cdot}95$ is in fact found from figure 4 to be $\tan^{-1}(-1{\cdot}8)$. The mean slopes for $\beta = 2{\cdot}56$ and $3{\cdot}598$ which are outside the range of figure 4 are still closer to the approximate value $\tan^{-1}(-2)$.

The velocity distribution is shown on a larger scale in figure 5 for the case when $\alpha = 0{\cdot}7$, $\beta = 0{\cdot}761$. The calculated points are marked in figure 5. The calculation has been carried beyond the point where the shock wave occurs and the corresponding part of the velocity distribution curve is marked in figure 5 with a broken line. It will be seen that in the virtual part of the curve the velocity for a given radius is no longer single-valued.

PRESSURE DISTRIBUTION

The pressure p at any point is given by

$$\frac{p}{p_0} = \frac{y}{y'}\left(\frac{\eta r^2\alpha^2}{R^2}\right)^{\gamma/(\gamma-1)}. \tag{39}$$

Values of p/p_0 for a selection of values of r/at are given in table 2. The pressure distributions for $\alpha = 0{\cdot}2$, $0{\cdot}4$, $0{\cdot}5$, $0{\cdot}6$, $0{\cdot}7$, $0{\cdot}8$, $1{\cdot}0$ and $1{\cdot}2$ are shown in figure 6. It will be noticed that those for $\alpha = 1{\cdot}0$ and $1{\cdot}2$ appear to be nearly parabolic. It can be shown in fact that the distribution is parabolic near the sphere so that when the thickness of the layer of expanding air is small compared with the radius of the sphere the distribution is nearly parabolic through the layer. If

$$s = (r-R)/r \quad \text{so that} \quad x = U_2(1+s), \tag{40}$$

and s is supposed small the approximate linear distribution of velocity is

$$u = U_2(1-2s),$$

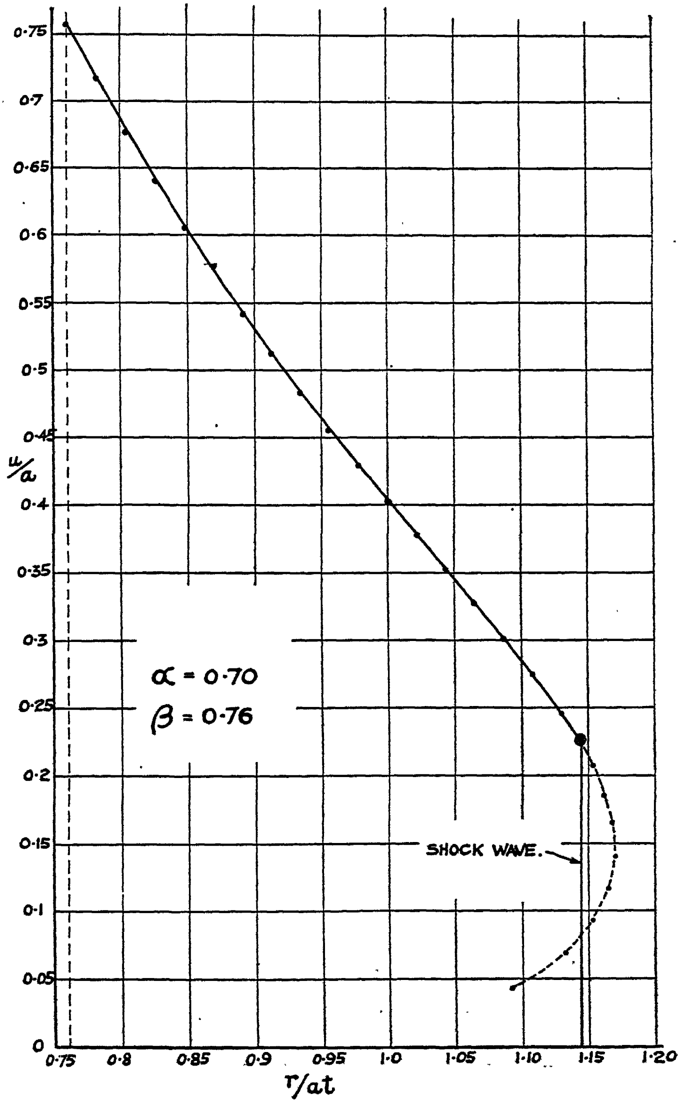


FIGURE 5. Distribution of velocity for $\alpha = 0{\cdot}7$.

equation (15) therefore takes the form

$$\frac{dc^2}{ds} = -6(\gamma-1)\,U_2^2 s,$$

so that $\qquad c_2^2 - c^2 = 3(\gamma-1)\,U_2^2 s^2 \quad \text{or} \quad 1-\frac{c^2}{c_2^2} = 3(\gamma-1)\,\alpha^2 s^2. \qquad (41)$

If s is small
$$\frac{p_2-p}{p_2} = \frac{\gamma}{\gamma-1}\frac{c_2^2-c^2}{c_2^2}$$

so that
$$1-\frac{p}{p_2} = 3\gamma\alpha^2 s^2 = 3\gamma\alpha^2\left(\frac{r-R}{R}\right)^2, \tag{42}$$

which represents a parabolic pressure distribution.

As an example of the application of the approximate expression (42) the values of y' and α corresponding with infinite rate of expansion may be

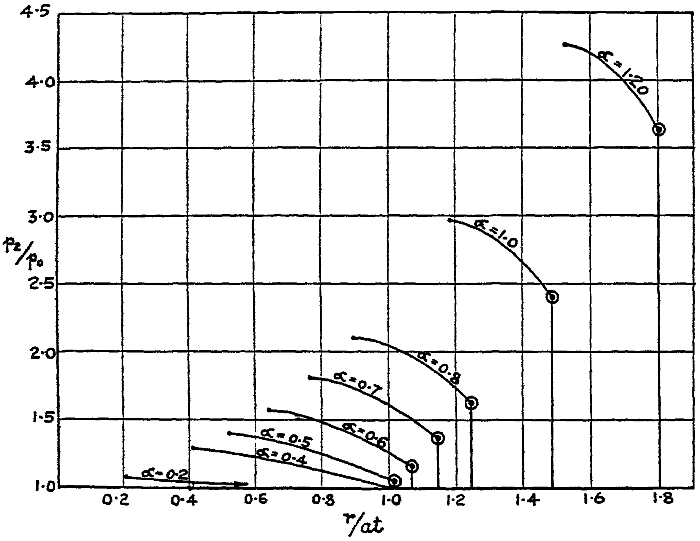


FIGURE 6. Distribution of pressure.

calculated. In this case $\xi_1 = 0{\cdot}8316$, $\eta_1 = 0{\cdot}1968$ (see equation (37)) so that s_1, the value of s at the shock wave is given by

$$\xi_1 = ut/r = 1-3s_1 = 0{\cdot}8316 \quad \text{and} \quad \eta_1 = c^2t^2/r^2 = \alpha^{-2}(1-2s_1).$$

Hence $s_1 = 0{\cdot}056$, so that

$$r_1/R = 1{\cdot}056 \quad \text{and} \quad \eta_1 = 0{\cdot}1968 = \alpha^{-2}\{1-2(0{\cdot}056)\},$$

so that $\alpha = 2{\cdot}12$. The approximate value of y' is $1-3\gamma\alpha^2 s^2$ or $0{\cdot}94$. These values may be compared with those given in table 1 which were found by numerical integration of the full equations, namely, $r_1/R = 1{\cdot}060$, $\alpha = 2{\cdot}103$, $y' = 0{\cdot}93$.

COMPARISON WITH SOUND WAVE SOLUTION FOR LOW RATES OF EXPANSION

For small values of α the sound wave solution of equations (8) and (9) may be expected to afford a good approximation to the true motion. The velocity and pressure distributions calculated from the complete equations and from the approximate equations of the theory of sound are compared in figure 7 for the case $\alpha = 0{\cdot}2$. So far as the velocity distribution is

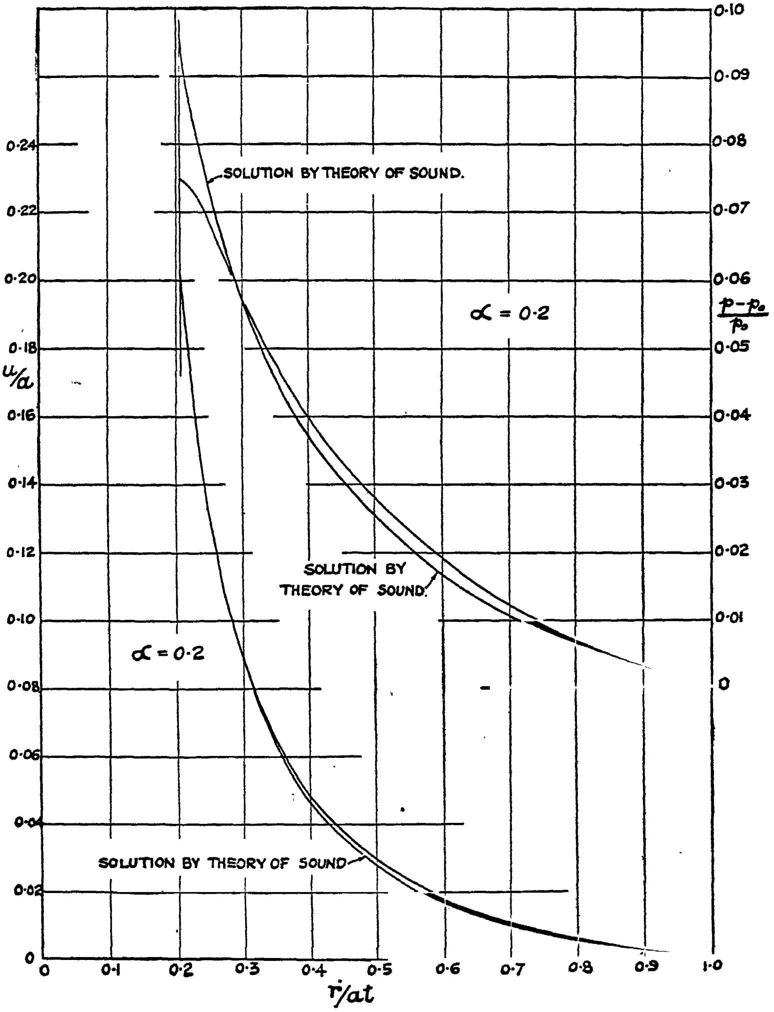


FIGURE 7. Comparison between velocity and pressure distributions calculated by the theory of sound and by the complete equations.

concerned the agreement is good but the pressure distributions are distinctly different near the sphere. The true pressure distribution is in fact parabolic near the sphere and initially $dp/dr = 0$. According to the sound wave equation (9) the value of dp/dr close to the sphere is $-\frac{2\rho\alpha^2}{1-\alpha^2}\left(\frac{a^2}{R}\right)$. The reason for the discrepancy is evidently that it is not justifiable to apply the equations of the theory of sound in the neighbourhood of the sphere owing to the neglect of the term $u\,\partial u/\partial r$ in the equation of motion in comparison with $\partial u/\partial t$. In the correct equation $u\,\partial u/\partial x$ is equal to $-\partial u/\partial t$ at the sphere.

Apart from this difference at the inner boundary the chief contrast between the sound wave solution and the true solution for values of α greater than 0·5 lies in the fact that the former involves no shock wave at the outer boundary. The true solution for $\alpha = 0{\cdot}2$ appears in figure 7 to resemble the sound wave solution in this respect. If this resemblance were true then some limiting value of α would exist below which no shock wave would be produced. Assuming that such a limit exists the author's rough attempts to determine its value placed it between $\alpha = 0{\cdot}4$ and $\alpha = 0{\cdot}5$. The matter was, however, examined later by Dr J. W. Maccoll using more accurate methods of numerical solution, and he came to the conclusion not only that a shock wave is formed when α is less than 0·4 but that no lower limit of the assumed type would be found. Subsequent analysis shows that this prediction is correct.

FORM OF SOLUTION NEAR $(\xi = 0,\ \eta = 1)$

Near the point $(\xi = 0, \eta = 1)$ equation (19) takes the form

$$\frac{d\zeta}{d\xi} = \frac{\zeta + (\gamma+1)\xi}{\xi}, \tag{43}$$

where $\zeta = \eta - 1$. The solution of equation (43) is

$$\zeta = (\gamma+1)\,\xi \log(A\xi), \tag{44}$$

where A is the constant of integration. As $\xi \to 0$, $\zeta \to 0$ but is negative when $\xi < 1/A$. Multiplying both sides of (44) by A it will be seen that $A\zeta$ is a function of $A\xi$ so that the shape of the (ξ, ζ) curve does not depend on the constant of integration though its scale is proportional to $1/A$.

In the neighbourhood of $(\xi_1 = 0, \eta_1 = 1)$ the relationship between ξ_1 and η_1 at a shock wave is found by taking $y - 1$ as small in equations (26) and (31).

Hence
$$\operatorname*{Lt}_{\xi_1 \to 0} \frac{\eta_1 - 1}{\xi_1} = -\frac{3-\gamma}{2} = -0{\cdot}797. \tag{45}$$

Comparing equations (45) with (44) the shock wave condition is satisfied if

$$\log_e A\xi = -\frac{3-\gamma}{2(\gamma+1)} = -0{\cdot}3316. \tag{46}$$

Hence

$$\left.\begin{aligned} A\xi &= +0{\cdot}718, \\ A\zeta &= A(\eta-1) = -0{\cdot}572. \end{aligned}\right\} \tag{47}$$

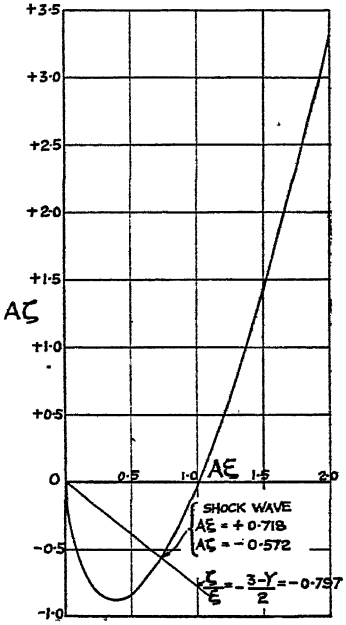


FIGURE 8

Figure 8 shows the form of the solution near ($\xi = 0$, $\eta = 1$) and the shock wave line intersecting the (ξ, η) curve at the point $(+0{\cdot}718,\ -0{\cdot}572)$. It may be concluded that provided the numerical solution brings the $(\xi\eta)$ curve into the neighbourhood of the point ($\xi = 0$, $\eta = 1$) the expanding air must be bounded by a shock wave. This indeed proves to be the case, but, as will be seen later, the shock wave is of very small intensity when α is less than 0·5.

FITTING THE NUMERICAL SOLUTION TO THE SOLUTION VALID NEAR $\xi = 0$

Writing $\zeta = \eta - 1$ equation (19) becomes

$$\frac{d\zeta}{d\xi} = \frac{2(1+\zeta)\zeta + (\gamma+1)\xi - \gamma\xi^2}{\xi} \, \frac{}{2+3\zeta+2\xi-\xi^2}. \tag{48}$$

In the approximate solution (44) ζ/ξ decreases as ξ decreases. In fitting the approximate solution to the numerical solution a pair of values of ξ and ζ may be taken and the value of A found by inserting these in equation (44). To find the error committed in using equation (43) instead of (48) suppose that the solutions are joined where $\zeta/\xi = B$ the true value of $d\zeta/d\xi$ at this point is

$$\frac{d\zeta}{d\xi} = \frac{(1+B\xi)(B+\gamma+1-\gamma\xi)}{2+(3B+2)\xi-\xi^2},$$

expanding this expression in powers of ξ the first two terms are

$$\left[\frac{d\zeta}{d\xi}\right]_B = (B+\gamma+1)\left\{1-\left(1+\tfrac{1}{2}B+\frac{\gamma}{B+\gamma+1}\right)\right\}.$$

The value of $[d\zeta/d\xi]_B$ in the approximate solution is $B+\gamma+1$. The proportional error is therefore

$$\epsilon = \left(1+\tfrac{1}{2}B+\frac{\gamma}{B+\gamma+1}\right)\xi$$

and since in the approximate solution

$$B = (\gamma+1)\log_e(A\xi),$$

the error in $d\zeta/d\xi$ is given by

$$A\epsilon = \left(1+\tfrac{1}{2}B+\frac{\gamma}{B+\gamma+1}\right)e^{(\gamma+1)B}. \tag{49}$$

Values of $A\epsilon$ for a series of values of B are given in table 3.

TABLE 3

B	-1	0	1	2	3	4	5	6	7	10
$A\epsilon$	0·96	1·58	2·9	5·5	9·6	16·9	29·4	50·2	84·4	387·0

$\alpha = 0·4$. When $\alpha = 0·4$ values of ξ and ζ given in lines 1 and 2 of table 4 were calculated numerically. Values of A calculated from equation (44) are given in line 3, and the proportional error calculated from equation (49) in line 5. The error ϵ is less than 5 % in the first three points.

TABLE 4. $\alpha = 0·4$

ξ	0·0081	0·0122	0·160	0·0231	0·0306	0·0384	0·0446	0·0536	0·0651
ζ	0·020	0·044	0·065	0·115	0·170	0·226	0·287	0·351	—
A	342	367	337	342	327	301	325	248	227
B	2·46	3·61	4·0	5·3	5·5	5·9	6·4	6·3	6·5
ϵ	0·021	0·038	0·046	0·105	0·12	0·16	0·19	0·24	0·29

Taking $A = 340$ the (ξ, η) curve shown in figure 9 was calculated. The first three values of η, namely, those corresponding with $\xi = 0{\cdot}0081, 0{\cdot}0122$ and $0{\cdot}160$ are shown in figure 9. It will be seen that though the curve calculated using equation (44) passes very nearly through these three points the existence of a portion of the $(\xi\eta)$ curve for which ζ is negative would not have been suspected from simple inspection of the apparent trend of the curve calculated step by step through 99 % of the range $\xi = 1$ to 0. This point may perhaps be appreciated more clearly if the area covered by figure 9 is compared with the same area (marked with broken line) on the much smaller scale of figure 1.

It appears from equation (47) that when $\alpha = 0{\cdot}4$ the expanding air is bounded by a shock wave for which

$$\xi_1 = 0{\cdot}718/340 = 0{\cdot}0021,$$

$$\eta_1 - 1 = -0{\cdot}572/340 = -0{\cdot}0017.$$

The corresponding change in pressure at the shock wave is $y - 1 = \gamma\xi_1 = 0{\cdot}003$ of an atmosphere.

$\alpha = 0{\cdot}5$. When $\alpha = 0{\cdot}5$ the four lowest values of η calculated numerically give $A = 24$, 21, 21 and 21. Taking $A = 21$ the shock wave corresponds with

$$\xi_1 = 0{\cdot}718/21 = 0{\cdot}034,$$

$$\eta_1 - 1 = -0{\cdot}572/21 = -0{\cdot}027,$$

so that $$y - 1 = 0{\cdot}034\gamma = 0{\cdot}048,$$

which agrees well with the value $y = 1{\cdot}05$ determined graphically by means of figure 1.

$\alpha = 0{\cdot}2$. When $\alpha = 0{\cdot}2$ the lowest values of ζ calculated numerically are

ξ	0·0002	0·0003	0·0005
ζ	+0·0175	+0·026	+0·043

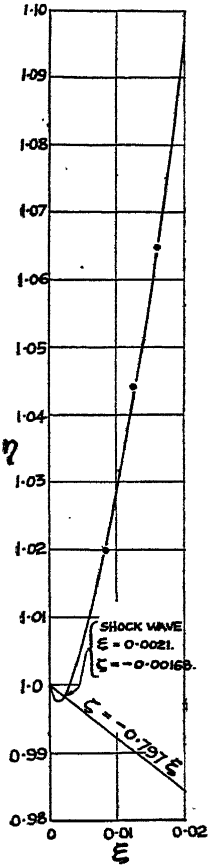


FIGURE 9. ξ, η curve near $\xi = 0$, $\eta = 1$, when $\alpha = 0{\cdot}4$.

When inserted in equation (44) these give values of ξ and η at the shock wave of order 10^{-19}. So far as the equations of a non-viscous fluid are concerned this shock wave seems to be real enough in spite of its extreme smallness but from the physical point of view such a minute shock wave has no meaning. The effect of viscosity and conductivity would in fact become appreciable long before a shock wave with pressure change 10^{-19} atm. could be formed. Nevertheless it is curious that there is this definite mathematical difference between a wave of finite intensity and the equivalent sound wave. It is especially curious that the point where the solution of the complete equation differs from the approximate solution of the theory of sound is in the region of very small velocities and pressure changes, the region in fact where the theory of sound might be expected to be most accurate.

In conclusion I wish to express my thanks to Mrs H. Glauert who carried out some of the calculations and to Dr J. W. MacColl for some valuable suggestions.

SUMMARY

The only case in which the motion of a gas at high speed in three dimensions has so far been discussed mathematically is that of the disturbance produced by a cone moving with velocity greater than that of sound. In the present work another case is analysed, namely, the radial outward flow produced by a uniformly expanding sphere. The region of expanding air is bounded by a shock wave outside which the air is undisturbed. As the radial velocity of the sphere increases the thickness of the layer of disturbed air decreases till at infinite rate of expansion it is only 6 % of the radius of the sphere. The distributions of velocity and pressure are given for a range of rates of expansion. When the radial velocity of the sphere is small an approximate analysis based on the theory of sound yields results which are inaccurate near the sphere and also at the shock wave which forms the outer boundary of the expanding air.

REFERENCES

Hugoniot, 1889 *J. Éc. polyt. Paris.*
Rankine, 1870 *Phil. Trans.* **160, 277.**

Explosion waves and shock waves

VI. The disturbance produced by bursting diaphragms with compressed air

By William Payman, D.Sc., Ph.D.
and Wilfred Charles Furness Shepherd, B.Sc., Ph.D.

(*Communicated by Sir Jocelyn Thorpe, F.R.S.—Received* 21 *March* 1940)

[Plates 15–18]

Experimental verification of the theoretical relationships governing the motion of shock waves has been derived from an investigation into the development of the disturbance set up in a uniform tube when a body of compressed gas, confined at one end by means of a copper diaphragm, is released by rupture of the diaphragm. The wave-speed camera has been used to obtain continuous Schlieren records of the passage of these effects along the tube and to analyse their properties under a variety of experimental conditions, the main variations being those caused by the use of different thicknesses of diaphragm, different lengths of compression and expansion chambers, and of different gases in these chambers.

The pressure effects released on the rupture of the diaphragm are the shock wave, a vortex formation and the expanding gases from behind the diaphragm. Schlieren snapshot (spark) photographs have been obtained showing the structure of these pressure effects when they are projected from the end of the tube into free air. The photographs show, in addition to these details, the establishment of a system of stationary sound waves.

Diagrams have been constructed based on the experimental data showing the apparent form of the wave during various phases of its progress.

It is well known that slight disturbances are communicated through the atmosphere at the normal speed of sound without giving rise to any sharp discontinuity, and the conditions of propagation of such disturbances have been examined in great detail both mathematically and experimentally. That more abrupt or more violent disturbances would tend to give rise to discontinuity, and would be propagated at speeds which could be much greater than that of sound, appears to have been suggested first by Stokes (1848) in discussing some earlier work by Poisson (1808). The mathematical treatment of the propagation of disturbances of this character has since been extended in great detail by Airy (1849), Earnshaw (1859, 1860), Rankine (1870), Schuster (1893), Chapman (1899), Rayleigh (1910) and Taylor (1910) in this country, by Riemann (1860), Templen (1905) and by Becker (1917, 1921, 1922) in Germany, and independently in France by Hugoniot (1887), Hadamard (1903), Duhem (1909) and Jouguet (1902, 1917). On the basis of their work, equations have been developed by means of which the properties of a shock wave travelling through a homogeneous medium can be calculated provided that one characteristic is known. Experimental investigations have been mainly concerned with shock waves produced by the detonation either of a gas mixture or of a solid explosive, in which circumstances they are accompanied initially by flame, gaseous products of combustion and solid particles, all of which influence the spread of the waves. A shock wave in a simpler form is produced when a diaphragm ruptures under a compressed-air load, the characteristics of the wave produced under these known conditions being determined solely by the volume and pressure of the air so released. The bursting diaphragm method was first used by Vieille (1899) to examine the propagation along a cylindrical tube of waves so produced, with the aid of a mechanical method for recording, at fixed points, the passage of the shock wave and gases. We have been able to apply the wave-speed camera (Payman and Woodhead 1931), which embodies the principle of Schlieren photography, in a more detailed analysis of the movement of such shock waves along a tube, and this investigation provides a basis for an experimental verification of the theoretical treatment.

APPARATUS

The apparatus used, in figure 1, consisted of a compression chamber C, and an expansion chamber E, both 2·54 cm. in internal diameter. The compression chamber was provided with end-flanges, to one of which was bolted a plate carrying two electrodes. The diaphragm, D, was clamped between the other flange and one end of the expansion chamber. The full length of the compression chamber was 61 cm., but its effective length was reduced when required in the following manner. A metal piston, 15 cm. long, with a fine hole drilled axially through it, fitted smoothly within the bore, and could be fixed at any point by means of set screws. This divided the compression chamber into two compartments, the one next to the

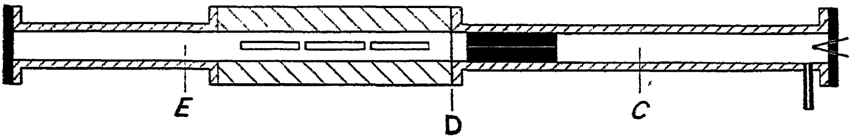


FIGURE 1. Diagram of apparatus.

diaphragm being the effective compression chamber. The flow of air through the fine axial hole in the piston when a diaphragm burst was too slow to have a material effect on the development of the disturbance. Compressed air was admitted through a union on the side of the chamber.

The expansion chamber consisted wholly or in part of the special steel vessel described in an earlier paper (Payman and Titman 1935). This vessel was 37 cm. long and was fitted with three windows along each side, so placed that a beam of light could pass through them, and Schlieren records could be obtained of the movement of the disturbance travelling along this portion of the expansion chamber. A longer expansion chamber could be obtained by inserting a length of tubing, of the same internal diameter, at either end of the steel vessel, and in this way different portions of the expansion chamber could be photographed.

During an experiment compressed air was slowly admitted to chamber *C* until the diaphragm ruptured. A disk of copper was sheared cleanly from the diaphragm and travelled along the expansion tube unless it was caught, as in some experiments, by means of a trap. This trap was a small piece of brass of aerofoil cross-section, mounted across the expansion chamber. The distance of the trap from the diaphragm was such that when the diaphragm was bulged outwards by the static air pressure it just failed to touch and be supported by the trap. When projected, the disk wrapped itself round the aerofoil and then offered no serious obstruction to the continued flow of air from the compression chamber.

The time of exposure of the photographic film was synchronized with the instant of rupture of the diaphragm by means of the shutter of the wave-speed camera. This shutter, as it fell in guides, uncovered a slit in front of the film for a period less than the time of one revolution of the camera drum. At the rear end of the compression chamber was placed a match-head, fired electrically, around which a few strands of gun-cotton were wrapped. Compressed air was passed into the compression chamber until the pressure attained was just below that necessary to burst the diaphragm. When the shutter was allowed to fall, it closed an electric circuit and fired the match-head at a predetermined position before the slit was uncovered. The additional pressure generated by the combustion of the match-head and gun-cotton was small but was sufficient to cause almost immediate rupture of the diaphragm, and the pressure effects passed the window in the expansion chamber during the interval of time available. Records show that the fumes from the gun-cotton did not spread into the expansion chamber for an appreciable time, and the pressure effects under examination were not modified by them in any way.

Bursting pressures of copper diaphragms. The diaphragms were cut from copper foil supplied in a convenient range of standard thicknesses. It was found that the bursting pressures of diaphragms cut from the same roll seldom varied more than ± 3 % from the average for that roll, and a similar variation was found in the average bursting pressures of diaphragms from rolls of the same nominal thickness purchased at different times. The average bursting pressures of diaphragms from various foils obtained over a period of years is given in table 1. The area of foil exposed to pressure was circular and 2·54 cm. in diameter. The thickness of the copper was within ± 0·0001 in. of the stated thickness for the thinner foils and within ± 0·0003 in. for the thicker foils. It is seen that the bursting pressure is roughly proportional to the thickness. A similar relationship was found by Bonyun (1935), but his absolute values were approximately double those we obtained. This marked difference may be due to the fact that he used copper of high chemical purity annealed after rolling in a non-oxidizing furnace, whereas the foils used in the present experiments were cold-rolled and not annealed. The chemical and physical properties of a diaphragm are therefore of importance in determining its bursting pressure. Another important factor is the nature of the 'cutting edge', that is, the rim of the flange around which the diaphragm ultimately shears. Difference in the profile of this edge may also account to a certain extent for the different results obtained by Bonyun.

TABLE 1

Thickness of copper in.	Static bursting pressure lb./sq. in. above atmospheric
0·001	60
0·002	115
0·004	230
0·006	350
0·008	460
0·010	600
0·016	900

PHOTOGRAPHIC RECORDS

Four typical wave-speed records are reproduced in figures 2–5 (plate 15), and the line diagram, figure 2*a*, has been added to assist their detailed description. The width of each record represents the full length (37 cm.)

of the steel vessel. The total length of tube illuminated and forming the Schlieren field of the wave-speed camera is only 29·1 cm., and the portion of the tube at each end which is not illuminated is shown as a broad vertical black band. The experimental arrangement of the tube for these records is as shown in figure 1, the diaphragm being situated at the right-hand end of the chamber. The record proper begins at a distance of 3·5 cm. from the diaphragm, this being the distance represented by the thickness of the black band. In all records the disturbance on its outward journey is shown as moving from right to left, and in the descriptions the windows will be numbered from right to left, the first window being the one nearest the diaphragm. The two narrower bands running vertically down each record are due to the interception of light by the window frames. In a similar way any opaque body, for example the copper disk, passing across the field of view, will intercept light and be recorded as an inclined black band. A compression wave moving from right to left is shown on the record as a white line; one moving from left to right, for example a reflected wave, is shown as a black line.

The propagation of the disturbance. Before describing the records it will be an advantage to consider how the disturbance is propagated within the system. Figure 6 shows diagrammatically the state of affairs inside the apparatus (*a*) before the diaphragm ruptures and (*b*) at the instant before the main shock wave reaches the end *B* of the expansion chamber. In

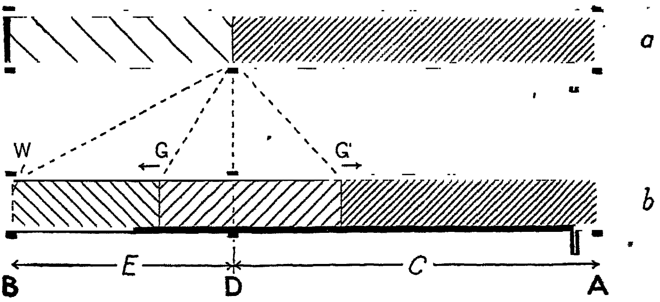


FIGURE 6. The spread of the disturbance in the compression and expansion chambers.

figure 6*a*, *DB* represents the column of air at atmospheric pressure in the expansion chamber, and *DA* the column of compressed air at room temperature in the compression chamber, the diaphragm being situated at *D*. When the diaphragm bursts, a shock wave *W* travels along the expansion chamber at a speed of 2–3 times that of sound. When the wave reaches the position shown in figure 6*b* some of the air in the compression chamber will have moved into the expansion chamber and its boundary will have reached a point G. Hence, the column of air DB originally filling the expansion chamber now occupies the column GB. We have called this volume of gas, between the shock wave and the gases creating the disturbance, the 'pressure wave' (Payman and Robinson 1926). At any instant during its outward journey the wave, W, is the boundary of all effects arising from the expansion. In the same way any effect can only be felt for a certain, shorter, distance along the compression chamber. The wave here, G', continuously marking the limit of the disturbance along the compression chamber, is a rarefaction wave, and its speed does not exceed that of sound. Hence, at the instant depicted, only the column GG' has begun to expand (and has cooled in consequence), the column $G'A$ being still at its original pressure.

Description of records. For the record shown in figure 2 the compression chamber was 61 cm. long and the expansion chamber consisted of the steel vessel only. The end of the expansion chamber remote from the diaphragm was closed by a metal plug. A copper diaphragm 0·002 in. thick was ruptured by an air pressure of 115 lb./sq. in. and the sheared disk was not trapped.

The first pressure effect from the bursting diaphragm is the shock wave W, shown as a thin white line. Immediately after this wave there emerges a series of faint traces (marked V in figure 2*a*), the nature of which will be discussed later. This is in turn followed by the front of the expanding air from the compression chamber, shown as a white band, and the copper disk, shown as a wide black band.

The shock wave W travels along the first window at a slowly increasing rate, accelerates more rapidly past the second window, and then proceeds at an approximately uniform speed to the end of the tube. Since the shock wave W is the pioneering effect of the disturbance resulting from the release of the compressed air, its increase in speed can only be due to some effect acting from behind. An examination of a number of the photographic records reveals that the acceleration is caused by a high-speed shock wave which overtakes and absorbs the primary wave. Though not visible in the reproduction of figure 2, its approximate position is shown by the dotted line W_1 in figure 2*a*. The feebler trace of the wave as it passes the central window is not due to a reduction in intensity of the wave. It is an optical effect caused by the shape of the wave. The wave is plane and does not cause appreciable refraction of the central portion of the beam of light. Displacement of the apparatus so that the centre of the beam passes through either of the end windows results in the wave being recorded feebly as it passes along that portion of the vessel.

When the shock wave W is reflected at the closed end B its speed relative to the wall of the tube is reduced. This, the 'apparent' speed, is not the true speed of the wave relative to the medium, which is in motion. The reflected wave, shown as a black trace, is marked RW in figure 2*a*; at O it meets the front of the air released from the compression chamber and undergoes a further loss in apparent speed.

The air escaping from the compression chamber travels at a fairly uniform speed, and its front is well defined (G, figure 2*a*) until it is met and retarded by the reflected shock wave. Prior to the collision the compressed air expands freely, but afterwards its expansion begins to be conditioned by the closed end B.

The width of the shadow cast by the copper disk gives an indication of the orientation of the disk with respect to the tube. When it first comes into view in figure 2 the disk is turning broadside to the windows and therefore edgeways to the air stream. When the band narrows later the disk has turned.

The fumes from the gun-cotton, which have not started to emerge from the compression chamber at the end of the time interval covered by figure 2, are shown at the lower right-hand corner of figure 3. For this record a diaphragm bursting at 430 lb./sq. in. was used, and the length of the expansion tube was increased to 67·7 cm. The steel vessel was next to the diaphragm. The reflected wave appears much later in this record, on account of the greater distance of travel beyond the portion photographed. The compression chamber was again 61 cm. long.

For the record reproduced in figure 4 a diaphragm bursting at 640 lb./sq. in. was used and the trap was fitted in the expansion chamber. The compression and expansion chambers were of the same length as for figure 3.

For figure 5 the disk was not trapped and the length of the compression chamber was reduced to 1·02 cm. The expansion chamber was again 67·7 cm. long, and the diaphragm burst at 430 lb./sq. in. as for figure 3. The effect of the one difference between these two experiments, the length of compression chamber, is well marked, for with the shorter chamber there is a more rapid decay in the speed of the air and the air expands to a much smaller distance.

The rupture of the diaphragm. These four typical diagrams help us to visualize the mode of rupture of the diaphragm, and to trace the origins of the various pressure effects. The sequence of events at the diaphragm cannot be observed directly, since the first 3·5 cm. of the tube are not photographed. The instant of departure of each pressure effect from the diaphragm, that is, from the end of the compression chamber, is estimated by producing its trace backwards, a procedure which may not be wholly free from error. This is done on the explanatory diagram (figure 2*a*) by the broken lines over the right-hand band. Zero time, at which the diaphragm ruptures, is the moment when the primary shock wave leaves the end of chamber *C*, for the traces of the wave and the copper disk unite at this point. Expansion of the compressed air is momentarily hindered by the disk of copper, and does not become comparatively free until the disk has travelled about a centimetre along the expansion tube and is turning edgewise to the air stream.

The shock wave. At the instant the diaphragm bursts a shock wave is sent along the expansion tube. This primary wave, *W*, is comparatively feeble and is followed by only a small volume of compressed air. A more powerful shock wave, W_1, is sent out later, when there is full release of the compressed air, and it soon overtakes and absorbs the primary wave. For the first 8 cm. or so of its travel, the stronger wave, which we may term the main wave, has an extremely high speed; later, constant speeds are maintained by both wave and compressed air, the expansion proceeding under more uniform conditions. The shock wave, the pressure wave and the expanding compressed air constitute the major pressure effects of the disturbance.

Vortex formation. There is, however, another pressure effect arising from the compression tube, namely, that giving rise to the traces marked *V* in the explanatory diagram. These traces are apparently caused by a small amount of compressed air which, escaping at the periphery of the sheared disk at the instant of rupture, follows the primary wave in advance of, and separated from, the main body of compressed air. It has been shown previously (Gawthrop, Perrott and Shepherd 1931) that on the emergence of a shock wave from the open end of a tube a vortex is set up in the rear of the wave, and this moves forward, maintaining its identity, at a comparatively high speed. The traces *V* probably arise from a similar formation set up in the expansion chamber, for it would appear, from the evidence of numerous experiments, that the air which first escapes from the compression chamber initially travels *through* the air in the expansion chamber as a discrete entity. Although there is little diffusion during the very early part of its travel, the vortex formation soon begins to disintegrate and the traces *V* to become irregular.

1. VARIATION IN BURSTING PRESSURES

We may now proceed to a discussion of the quantitative experiments in which the speeds of the pressure effects were measured with and without the trap designed to catch the sheared disk. The compression chamber was 61 cm. and the expansion chamber 67·7 cm. long, and end B was permanently closed by a brass plate.

(*a*) *Experiments without trap.* The photographic records are similar in type to that shown in figure 3, though at higher pressures they become more confused owing to the increased intensity and closer proximity of the pressure effects. In table 2 are recorded the speeds of the primary shock wave W, and the initial and final speeds of the main shock wave W_1. The table also gives the speeds of the compressed air G, and of the copper disk D, both of which are sensibly uniform. The speed of the vortex formation V is not given as, in general, it is the same as G. All measurements are made from the photographic records and speeds are relative to the walls of the tube. The photographic field begins 3·5 cm. from the diaphragm and extends for a further 29·1 cm. Each value is the mean of three or more repeat experiments.

TABLE 2. EFFECT OF DIFFERENT PRESSURES IN COMPRESSION CHAMBER. WITHOUT TRAP

Bursting pressure lb./sq. in. above atmospheric	Mean speeds, m./sec.				
	Primary shock wave	Main shock wave		Compressed air, over whole field	Copper disk, over whole field
		Over first 8 cm.	Final uniform speed		
60	420	—	495	225	125
90	440	—	540	275	160
120	455	—	560	295	160
170	470	—	600	340	165
225	480	900	610	370	190
300	495	—	—	410	190
360	510	1000	670	440	180
430	530	1000	700	445	200
530	540	1100	710	475	200
715	570	1100	745	520	175
920	585	1120	770	585	155
1100	595	1170	770	585	145

There is at first a steady increase in the speed of the primary and main shock waves as the bursting pressure increases, the lowest speed recorded being much higher than that of sound. The acceleration of the primary shock wave due to the arrival of the main shock wave is evident from a comparison of columns 2 and 4. It is the speed of the main wave which is the true index of the pressure level in the compression chamber. The results suggest that, as the bursting pressure is increased, the speed of the shock wave produced tends towards a maximum value, as does also the rate of flow of the compressed air. This is shown in figure 7, where the steady speeds of the main wave, *W*, and the speed of the compressed air, *G*, are plotted against the bursting pressures of the diaphragms.

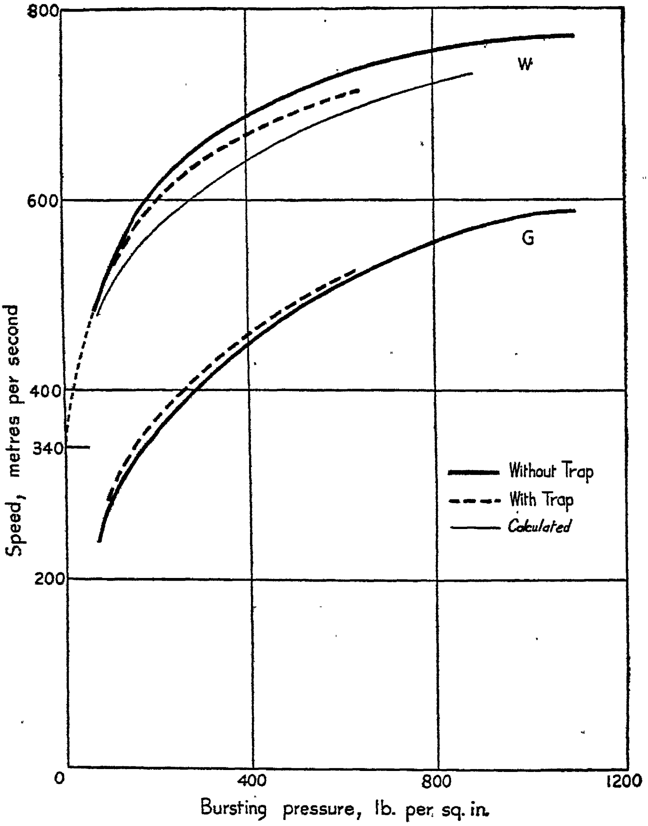


Figure 7. The effect of bursting pressure on the speed of the disturbance.

(*b*) *Experiments with trap.* The experimental arrangement was similar to that used without the trap, but a different 'cutting edge' altered slightly the bursting pressures with diaphragms of a given thickness. Speed measurements are given in table 3.

When these results are compared graphically with those given in table 2 (see figure 7), it is seen that the presence of the trap makes little difference either to the final speed of the shock wave or to the speed of the compressed air. The main difference introduced by the trap is in the early stages, immediately after the diaphragm has ruptured; the release of the compressed air appears to be more continuous, for the trap begins to function immediately the disk is released. Consequently the compressed air escapes at a progressively increasing rate as the copper disk rolls itself round the trap and the formation of a single powerful wave is avoided. Instead, there is a series of feeble waves which ultimately build up the primary wave into a main wave of intensity similar to that given by the same bursting pressure without the trap. The behaviour of the wave and gases is thereafter the same, so that the final uniform speeds are not sensibly altered by the presence of the trap.

TABLE 3. EFFECT OF DIFFERENT PRESSURES IN COMPRESSION CHAMBER. COPPER DISK TRAPPED

Bursting pressure lb./sq. in. above atmospheric	Mean speeds, m./sec.		
	Primary wave	Main wave, final uniform speed	Compressed air, over whole field
65	430	500	245
90	450	535	265
110	465	550	305
200	490	590	380
300	510	635	420
410	535	665	450
470	570	680	470
550	560	695	495
640	586	710	530

The reflected wave. The wave RW reflected from the closed end B is of importance in the study of the later effects arising from the compression chamber. In all the records of this series of experiments the front of the air from the compression chamber has passed all three windows some time before the reflected wave first appears in the left-hand window (at a time about 2·7 msec. in figures 3 and 4). Thus the reflected wave is now travelling through air from the compression chamber moving in the opposite direction and, until it passes into the gases from the gun-cotton, its apparent speed is approximately constant. This indicates that there is little change in the speed of the air current encountered; the continued steady rate of flow is important, for it remains comparable in value with the initial rate which was responsible for the uniform speed of the main shock wave. The apparent speed of the reflected wave after it enters the air from the compression chamber is approximately the same whatever the bursting pressure used (with or without the trap), and is 200–210 m./sec. Not only is the speed independent of the initial pressure in the compression chamber, but so also is the time of appearance of the reflected wave measured from the moment of rupture of the diaphragm. This means that the shock wave, on its journey to the closed end and back again to the starting point, travels at an average speed which is independent of the bursting pressure. Moreover, the time taken for the whole journey is almost exactly the same as the time taken by a sound wave to travel the same distance through air at N.T.P. This remarkable result is true also for all expansion systems used when chamber *C* was not shorter than chamber *E*. The time taken for a sound wave to travel the return journey at 340 m./sec. would be 3·8 msec. With pressures varying from 60 to 1100 lb./sq. in. the average time taken by the shock wave is 3·95 msec. When the trap is present the average time is 3·90 msec. With a compression tube 61 cm. long and an expansion tube 37 cm. long the times are 2·03 msec. for the sound wave and 2·10 msec. for the shock wave. Similar results were obtained with a still shorter expansion chamber.

This agreement indicates that there is a relationship between the true speed of the shock wave and the speed of the following air current. The proviso that the length of chamber *C* is similar to that of *E* merely ensures that the steady rate of flow from the compression chamber is maintained until the shock wave has returned to its starting point. Under these conditions the speed of the wave is so governed by the expansion of the compressed medium that, in so far as the outward journey is made in less time than would be required by a sound wave, then the time required for the return journey is correspondingly increased as a result of the changed physical condition of the medium encountered by the reflected wave.*

On impact with the closed end *B* of the expansion chamber the main shock wave is reflected and the pressure there is almost instantaneously increased. Travelling away from the closed end, the wave encounters and stems the oncoming stream of compressed air which hitherto has been moving forward at a uniform rate in the process of pressure equalization. From the moment of arrival of the wave, the pressure at the closed end will rise or fall, in diminishing surges due to repeated reflexion of the shock wave, until finally a steady pressure is reached, the level of which is determined by the bursting pressure and the lengths of compression and expansion chambers.

---

* Herein probably lies the explanation of a puzzling feature of records of flame propagation and detonation. Dixon, in his well-known experiments (1903), was able to explain the movements of flame on the assumption that a sound wave travelled in advance of the flame and then was reflected to meet the flame and retard it. It was shown later (Payman 1928) that the sound wave from the spark was of comparatively negligible intensity by the time it reached the end of the tube. The arrest was due to a powerful shock wave travelling at a speed much greater than sound and reflected in the same way. Yet the point of arrest was never far removed from that calculated on the assumption that the effect was due to a sound wave.

2. VARIATION IN LENGTH OF COMPRESSION CHAMBER

For the next series of experiments the compression chamber was shortened by means of the piston described earlier, the match head and gun-cotton being fired in the rear compartment. An expansion chamber 67·7 cm. long was used, with diaphragms bursting at 430 lb./sq. in. No difference was observed in the speed of the main shock wave when the length of the compression chamber was reduced from 61 to 6·35 cm., and the average speed at which the compressed air traversed the whole field was reduced only by a few metres per second. The results obtained without a trap are summarized in table 4; repeat experiments with the trap gave approximately the same results. Figure 5 shows the disturbance with a compression chamber 1·02 cm. long, and this should be compared with figure 3 in which the compression chamber was 61 cm. long.

TABLE 4. EFFECT OF LENGTH OF COMPRESSION CHAMBER

Length of compression chamber cm.	Speed m./sec.		
	Primary shock wave	Main shock wave, uniform speed	Compressed air, over whole field
61·0	530	700	445
6·35	530	700	420
2·54	535	665	310
1·02	520	600	Field not covered
0·42	500	520	,,

That increase in length of the compression chamber above about 6 cm. has no appreciable effect on the speeds of the main wave and of the front of the compressed air is due to the fact that the whole column of compressed air does not begin to expand immediately the diaphragm ruptures; the start of expansion spreads progressively layer by layer through the compression chamber. For example, if the whole column has not begun to expand when the main shock wave and the front of the compressed air have both passed out of the field of view (cf. the typical record, figure 3), the progress of these effects up to this stage is quite unaffected by the column of gas still fully compressed in chamber *C*. The length of the column of air which has begun to expand at any instant is defined by the distance which the rarefaction wave has travelled along the compression chamber, and records of this wave, obtained by using the vessel with windows as the *compression chamber*, show that it moves at approximately the speed of sound. The rarefaction wave is ultimately reflected at the closed end *A* and then follows the main wave, its apparent speed being increased because it is now moving in a medium which, though cooled by partial expansion, is itself moving in the same direction. With short compression chambers the influence of this rarefaction wave will soon result in a reduction, first in the rate of expansion of compressed air and, later, in the speed of the main wave. With longer compression chambers the degrading effect of the reflected rarefaction wave will be proportionately delayed. Thus the sole effect of increase in length of the compression chamber, that is, of increase in the volume of air released, is to extend the period during which the wave and compressed air are maintained at their steady speeds. The records obtained in these experiments show the characteristics previously described, these being the primary wave, the vortex formation and the main body of compressed air which gives rise to the main shock wave. The identities of the vortex formation and the main body of compressed air are very clearly brought out in figure 5, there being a perceptible interval of time between their appearances.

It will be seen from table 4 that the speed of the primary wave is independent of the length of the compression chamber, a finding which is in keeping with the suggested origin of this wave. On the release of the main body of the compressed air the main shock wave is set up at a speed which is also independent of the length of the compression chamber; this speed over the first 8 cm. travel is about 1000 m./sec. The final speed of this wave is, however, governed to some extent by the length of the compression chamber, since when this is reduced below 6·35 cm., the speed of the main wave is also reduced. With the shortest compression chamber, 0·42 cm., the primary wave is not overtaken by the main wave (which is rapidly degraded), and its speed remains practically unchanged.

3. VARIATIONS IN THE EXPANSION CHAMBER

It had been expected that the copper disk, if allowed to travel along with the several pressure effects, would have a large influence on their mode of propagation. Actually its influence was found to be small, for, as is seen from the records, the disk soon lags far behind. There is thus no need for the trap in the expansion chamber, and since the presence of the trap has been found to cause some confusion during the early stages of the disturbance, its use was discarded in the later experiments.

Conditions at the end of the expansion chamber. For figures 8 and 9, plate 16, the expansion chamber was 67·7 cm. long and the compression chamber 61 cm. long; diaphragms bursting at 430 lb./sq. in. were used. The expansion chamber was made up of the steel vessel and an extension tube, and by interchanging these two sections we could photograph the disturbance in each half of the expansion chamber. Figure 8 is a composite record showing the spread of the disturbance throughout the closed expansion chamber. For figure 9 the end of the expansion chamber was open, and the disturbance is shown spreading through the chamber and then outwards in free air for a distance of 90 cm.

A comparison of these two figures shows that the disturbance develops along identical lines until the main shock wave reaches the end of the expansion chamber. Thereafter, the development of the disturbance in the chamber is governed by the conditions at the end. In figure 8 the uniform speed of the main shock wave is 700 m./sec.; this speed persists until the wave is reflected at the closed end. The front of the air from the compression chamber retains its initial speed of 445 m./sec. until it collides with the reflected shock wave, whereupon the air is retarded and the speed of the wave reduced. At the time of the collision, the whole column of compressed air in chamber C has not started to expand, and the prevailing conditions of expansion correspond with a *compression* chamber of infinite length. The confining influence of the closed end at B is not felt by any moving layer of air until the layer is absorbed in the reflected wave RW, and until this occurs each layer continues to move as though the *expansion* chamber were of infinite length. The apparent speed of the shock wave after reflexion is 350 m./sec., and this speed falls to about 200 m./sec. soon after the wave enters the air released from the compression chamber. The uniform speed of wave RW (350 m./sec.) suggests uniformity of conditions in the air in the pressure wave which is, of course, heated by adiabatic compression, thereby causing the intrinsic speed of the wave RW to be *increased*; the further apparent loss in speed which the wave suffers on entering the air from chamber C is due largely to the reduced temperature.

In figure 9 the shock wave also maintains the speed of 700 m./sec. until it reaches the open end. The front of the air from the compression chamber travels at a uniform speed of 445 m./sec. until it is rapidly accelerated at a point within 3·5 cm. of the open end. This phase of accelerated motion is not shown on the record because it occurs in the chamber beyond the end window; it can be deduced, however, from the times at which the front of the compressed air disappears from view at the end of the last window and then reappears beyond the open end, and for a distance of 15 cm. beyond the end the air travels at a mean speed of 700 m./sec. The acceleration of the air front is due to the open end, and its cause has been discussed in an earlier paper (Payman and Shepherd 1937). When the main shock wave reaches the open end it begins to spread spherically in free air; its speed falls rapidly and is only 360 m./sec. at a point 30 cm. from the end. As it emerges from the tube the wave sets up a well-defined vortex ring which remains intact for some 30 cm. travel, over which distance it has an average speed of 215 m./sec. This vortex ring is a phenomenon quite distinct from the vortex formation set up within the expansion chamber at the instant the diaphragm ruptured. (As is clear from the dispersion of trace *V*, this latter formation is broken up before it reaches the open end.) The air from the chamber *C* appears to suffer a marked deceleration when it passes through the vortex ring; it preserves a well-defined front over a distance of at least 90 cm. at the end of which its speed is still 165 m./sec.

We have obtained spark photographs showing the configuration of the disturbance beyond the end of the expansion chamber. The distance scale on these records is 2·1 times that of the wave-speed records and the various traces are reversed, black to white and vice versa. Figures 10–13, plate 17, show early stages in the growth of the shock wave and of the accompanying vortex ring,* the experimental conditions being the same as those in the wave-speed record (figure 9). Clearly shown on figures 12 and 13 is the fine-line structure in the vortex ring and in the air stream, which thus far is composed solely of the air layers which constituted the pressure wave within the tube. Figures 14–17, plate 18, show the later stages in the disturbance and it is clear that there is little lateral expansion of the air stream; it would appear almost as though the air is constrained to travel in an imaginary cylinder bounded at the front by the vortex ring. The fine-line structure is a Schlieren effect set up by a stationary sound-wave system; one well-marked white trace persists throughout the records, and its counterpart is equally well marked on the wave-speed record (the trace being black in this case) at a distance of about 5–7 cm. from the end of the chamber. The wave causing this trace remains almost stationary throughout the whole history of the disturbance, although there is a slight

* The photographs may be compared with the records of vortex rings produced by water jets flowing into water obtained by Banerji and Barave (1931).

shift outwards when the compressed air from chamber C begins to emerge. This shift is shown by the wave-speed record and, on the snapshot record, it occurs between figures 14 and 16. On the intermediate record (figure 15), the air from chamber C has just started to issue from the expansion chamber. On figures 16 and 17 the traces of the shock wave apparently passing through the air stream are caused by the main shock wave which, having passed out of the normal field of view (defined by the reflected beam of light), is giving rise to a secondary Schlieren effect as it passes through the incident beam. The true positions of the wave in each of these two illustrations are 26 cm. ahead of the positions shown.

There is another interesting feature of all wave-speed records showing the discharge of pressure effects from the open end of an expansion chamber. Sooner or later, there is developed on the records a banded appearance. With low pressures, the bands are formed almost immediately the shock wave emerges into free air; with high pressures, they only become evident when the discharge has lasted several milliseconds. The banded appearance is most clearly brought out when a diaphragm bursting at a low pressure is used in conjunction with a low film speed, and figure 18, plate 15 was obtained in this way; it shows the external disturbance originated by a diaphragm bursting at 120 lb./sq. in. The bands begin to form before the gases originally contained in the compression chamber start to emerge, an occurrence which causes a sudden shift outwards of the first two or three bands. Then follows a period during which the bands, apart from a momentary reversal due to the copper disk, drift slowly towards the end of the expansion chamber until finally they converge rapidly.

These bands are the Schlieren traces of a series of 'stationary' sound waves in the stream of gases issuing from the expansion chamber, and the drifting of the bands is a natural corollary of fluctuations in the speed, temperature and pressure of these gases. The characteristics of such stationary waves, produced when jets of gas issue from small nozzles, have been examined by Emden (1899), Prandtl (1907), Stanton (1926) and Hartmann (1931). In our experiments, the gases issue, abruptly at first, from a uniform tube, 2·54 cm. in diameter. Figures 19 and 20, plate 18, are records obtained under the same experimental conditions as figure 18 and they show the criss-cross structure of the waves within the air stream which causes the banded appearance on the wave-speed records. (The shock wave shown on figure 20 is due to the secondary Schlieren effect already mentioned.) On many wave-speed records, there is evidence of a similar banded appearance in the air stream *within* the expansion chamber, the bands again being due to stationary waves. This effect is faintly shown, for example, in figure 3, plate 15, a system of about ten bands being visible just behind the front of the air from the compression chamber.

The vortex formation within the expansion chamber. We have suggested that the primary wave set up by the rupture of the diaphragm is accompanied by some kind of vortex formation which gives rise to the traces marked *V* in figure 2*a*, plate 15. This vortex formation, though apparently stable in the early stages, soon begins to suffer dispersion as is to be expected with any such system travelling along a uniform tube.

In our photographic studies of shock waves, it has been invariably found that a vortex of the simple ring type is formed at the mouth of any orifice through which a shock wave is made to pass. Vortices so formed are found to behave in much the same manner as do simple and slowly moving 'smoke-rings'. For example, the passage through an orifice of two shock waves in quick succession results in two vortices, the second of which will gather speed and, on entering the sphere of influence of the first vortex, will shrink to pass through it, an interchange of positions which is then repeated. Figures 10–13, plate 17, show the structure of this simple type of vortex which can retain its identity and clarity of outline despite a high rate of advance, in these illustrations 215 m./sec. By analogy with smoke rings, it is assumed that the vortex set up by a shock wave is composed of a discrete quantity of the gas originally *within* the orifice and that for some appreciable time there is little or no diffusion at the bounding surface as the ring expands.

We have used this property of a shock wave in a photographic study of the behaviour of vortices formed inside the expansion chamber. In one series of experiments the requisite orifice was arranged in a brass plate mounted across the expansion chamber not less than 30 cm. from the diaphragm. The shock wave travels along the expansion chamber and, on reaching the restriction, part is transmitted through the orifice and the remainder is reflected from the solid portion of the restriction. Schlieren records show that under these conditions a well-defined vortex is set up by the transmitted wave; the stability of the vortex depends upon numerous factors, in particular on the strength of the disturbance reaching the restriction and on the diameter of the orifice, although this was not smaller than 1·25 cm. Under favourable conditions, the vortex, recognizable by its distinctive trace, is found to travel at least 10 cm. before dispersion begins and traces still persist after a further similar distance. The records also suggest that the apparent high speed of vortices so formed is due largely to a motion of the medium itself and that their intrinsic speed relative to the medium is quite small. In a second series of experiments the conditions under which the vortex was formed were made to correspond more closely with those obtaining at the instant the copper disk is sheared from the diaphragm and the air disturbance is released. The restriction was a solid disk specially mounted across the expansion chamber so as to leave a narrow annular gap between the disk and the wall of the chamber. A small proportion of the incident wave is transmitted through this gap and gives rise to a forward motion of the gases in the gap, initially in the form of a vortex.

4. USE OF GASES OTHER THAN AIR

(*a*) *In compression chamber.* It has been stated that the main shock wave is maintained at a steady velocity by the movement of compressed air behind it, the velocities of the air and the wave being maxima for the existing conditions of temperature, initial pressure, etc. This would suggest that, other things being unchanged, a faster rate of flow from the compression chamber should result in a higher speed of shock wave through the air in the expansion chamber. This rate of flow could be increased by heating the air in the compression chamber but it was found difficult to confine the heating to the compression chamber alone. A simpler method of varying the flow was by altering the density of the gas used in the compression chamber, and in table 5 are recorded the results obtained with carbon dioxide and hydrogen. Diaphragms bursting at 120 lb./sq. in. were used and the expansion chamber contained air.

The photographic records obtained show the same features as with air in the compression chamber. The main shock wave and the front of the gases from the compression chamber both develop uniform rates. The effect of the change of the rate of outflow of the compressed gas is as was anticipated, that is, the higher the rate of flow of compressed gas the higher the speed of the shock wave. Since the shock wave is travelling through air in the expansion tube, its speed should be the same for a given rate of movement of gas from the compression chamber whatever the nature of

TABLE 5. DIFFERENT GASES IN COMPRESSION CHAMBER. AIR IN EXPANSION CHAMBER

Gas in compression chamber	Speed of primary wave m./sec.	Final speed of main wave m./sec.	Mean speed of compressed gas m./sec.
Carbon dioxide	420	520	270
Air	455	560	295
Hydrogen	605	735	Approx. 550

this gas, and this is shown to be true by a comparison with the corresponding values obtained by interpolation from the results given in table 2. The speed of air flow of 270 m./sec. gives rise to a shock wave with a final uniform speed of 530 m./sec., which agrees well with the carbon dioxide value. A speed of air flow of 550 m./sec. gives the corresponding value of 750 m./sec., which agrees equally well with the hydrogen value.

(*b*) *In expansion chamber.* Experiments have also been carried out with different media in the expansion chamber, air being used in the compression chamber. The diaphragms used burst at 225 lb./sq. in. The results are recorded in table 6, the speed of sound in the gas in the expansion chamber at 10° C being added for purposes of comparison.

TABLE 6. DIFFERENT MEDIA IN EXPANSION CHAMBER.
AIR IN COMPRESSION CHAMBER

Gas in expansion chamber	Final speed of main wave m./sec.	Speed of sound at 10° C m./sec.	Mean speed of compressed air m./sec.
Carbon dioxide	510	260	320
Air	610	340	370
Hydrogen	1480	1296	475

These experiments serve to show that the rate of propagation of the disturbance is determined by the nature of the medium in much the same way as is the speed of sound; the effect of the density of the gas in the expansion chamber on the rate of expansion of the compressed air is also well marked. The photographic records are similar to those already described except that with hydrogen no Schlieren trace of the wave is obtained, and the speed given in the table has been deduced from records obtained with an open-ended expansion tube, for the wave, though not recorded when passing through the hydrogen in this chamber, gives a well-marked trace when it reaches the outside air. It may be noted that the speed of the main shock wave is little above the speed of sound in hydrogen, and this accounts in part for the feeble optical effect produced by the wave.

THEORETICAL

The exhaustive mathematical treatments of the formation and propagation of shock waves have led to the development of equations which correlate the speed of a given wave with the temperature and pressure generated in the wave. For convenience we have adopted those given by

Becker (1922) though they are essentially similar to those given by earlier workers. The following equations apply to a plane shock wave propagating in a uniform tube as in our experiments:

$$\frac{T_2}{T_1} = \pi\frac{\pi+\phi}{\pi\phi+1}, \tag{1}$$

$$\frac{V_2}{V_1} = \frac{\rho_1}{\rho_2} = \frac{T_2}{\pi T_1}, \tag{2}$$

$$D^2 = P_1 V_1 \frac{\pi\phi+1}{\phi-1}, \tag{3}$$

$$U^2 = \left\{\frac{P_1V_1(\pi-1)}{D}\right\}^2, \tag{4}$$

where D = speed of shock wave,

U = speed of gases behind the shock wave,

T_2 = absolute temperature in the shock wave,

P_2 = absolute pressure in the shock wave,

$V_2 = 1/\rho_2$ = specific volume in the shock wave,

T_1 = absolute temperature in the undisturbed medium,

P_1 = absolute pressure in the undisturbed medium,

$V_1 = 1/\rho_1$ = specific volume of the undisturbed medium,

$\phi = 2\bar{C}_v/R + 1$, where $\bar{C}_v$ = mean specific heat, at constant volume, between T_1 and T_2,

$\pi = P_2/P_1$.

We have used these equations to calculate the average conditions of pressure and temperature in the main wave during the first 8 cm. of its travel and during its later uniform motion, using the measured speeds given in table 2 as the known characteristics. The values are given in table 7, it being assumed that the initial temperature was 0° C and that the initial pressure in the expansion chamber was 1 atm. (14·7 lb./sq. in.).

Pressure. It will be seen that the calculated average pressure in the shock wave during the initial stage is always considerably less than the pressure in the compression chamber when the diaphragm bursts, being about half at a bursting pressure of 16·3 atm. and only one-fifth at a bursting pressure of 75·8 atm. It must be remembered that these are average pressures in the wave during a phase of its existence in which it is rapidly degenerating, and that the initial instantaneous values must be considerably higher. Although at the moment the diaphragm bursts, the pressure discontinuity in the wave may be of the same order as the bursting pressure, more nearly approaching it the weaker the disturbance, it is not to be expected that the pressure can rise as high as that in the compression chamber for the speed of the wave, D, and hence also the pressure, P_2/P_1, are decided by the speed, U, of the gases moving forward behind it. The compressed air is at this moment stationary (that is $U = 0$), and since it can move forward only with a loss in pressure, it is impossible for the wave to be propagated for any appreciable distance with characteristics associated with the initial pressure discontinuity between the compression and expansion chambers.

TABLE 7. PROPERTIES OF MAIN WAVE

Bursting pressure		Initial stage. Average conditions during first 8 cm. of travel			Final stage of steady conditions during uniform propagation		
lb./sq. in. above atm.	Atm. abs.	Speed m./sec. Measured	Pressure atm. abs. Calc.	Temp. °C Calc.	Speed m./sec. Measured	Pressure atm. abs. Calc.	Temp. °C Calc.
225	16·3	900	8·5	370	610	3·8	157
360	25·5	1000	10·4	450	670	4·6	190
430	30·2	1000	10·4	450	700	5·0	209
530	37·0	1100	13·0	550	710	5·2	217
715	49·6	1100	13·0	550	745	5·7	244
920	63·6	1120	13·5	570	770	6·2	263
1100	75·8	1170	15·0	630	770	6·2	263

The movement of gases behind the wave. The theory upon which the equations used in these calculations are based requires that the shock wave must be accompanied by a movement in the same direction of the gases through which the shock wave has propagated and that a particular speed D of the shock wave is always associated with a particular speed U of the gases in its immediate rear. Since, in the present experiments the speeds of both the wave and the following gases have been measured, the results can be used to check the accuracy of the assumption. In the following comparison the results with short compression chambers are not applicable, for with a short chamber the main wave is followed too soon by the rarefaction wave from the emptying compression chamber and steady conditions are not attained.

It is immaterial whether we regard the shock wave as the cause of the gas movement set up in its rear or as being maintained and its speed determined by the gas flow in its rear. The photographic records show that the gases in the pressure wave are moving forward at a uniform speed, for the reflected wave travels through them at constant speed. Moreover, the speed of the gas layers in the pressure wave must be the same as that of the front layers of the gases from the compression chambers so that the latter speed may therefore be taken as corresponding with the value U required by the theory. In table 8 the speed of gases, U, calculated for the measured speeds, D, of the wave are compared with the actual rates of air flow as determined experimentally.

TABLE 8. RATE OF FLOW BEHIND THE WAVE FRONT.
AIR IN COMPRESSION AND EXPANSION CHAMBERS

Measured speed D of shock wave m./sec.	Measured speed of gases from compression chamber m./sec.	Calc. speed U of gases behind shock wave of speed D m./sec.
495	225	227
540	275	281
560	295	304
600	340	349
610	370	360
630	410	380
670	440	421
700	445	452
710	475	458
745	520	498
770	585	525

Except for the last value the agreement of the calculated with the observed results is remarkably close.

A further comparison arises from results of the experiments in which carbon dioxide and hydrogen were used in the compression chamber in place of air (table 7). Air was used in the expansion chamber and it was found that the speed of the shock wave therein depends upon the speed of the gases from the compression chamber, being somewhat slower when these gases are carbon dioxide and appreciably higher when hydrogen is used in the compression chamber. The front layers of the carbon dioxide travelled outwards at a speed of 270 m./sec., those of the hydrogen at a speed of 530 m./sec., giving rise to shock waves of speeds of 520 and 735 m./sec. respectively. The calculated values of U, corresponding to these speeds of shock waves are 260 and 490 m./sec. respectively, and the close agreement indicates that the speed of the shock wave, passing in each instance through air, depends solely upon the speed of the maintaining gases.

The form of the wave. Although a shock wave is generally referred to as a discontinuous wave, it has been shown by Rayleigh (1910) and Taylor (1910) that the thickness of the wave is finite. Nevertheless the thickness is extremely minute, being only 66×10^{-7} cm. for a wave travelling through air at a speed of 1000 m./sec. (The mean free path in air is about 90×10^{-7} cm.) The thickness of the front of the wave produced in different gases by a diaphragm bursting under a compressed-air load of 225 lb./sq. in. (table 6) is 88×10^{-7} cm. in carbon dioxide, 159×10^{-7} cm. in air and 1700×10^{-7} cm. in hydrogen. The comparatively high value for hydrogen is significant. It will be recalled that in these experiments the speed of the wave in hydrogen was not greatly above that of sound and that the wave produced no Schlieren effect; hence it may be concluded that the establishment of a discontinuity in any medium becomes increasingly difficult the higher the speed of sound in that medium and that the slope of the wave front is more drawn out the less dense the medium. The latter conclusion is supported by the calculations cited above, the wave front in hydrogen being more than 10 times as deep as that in air. Furthermore, the pressure in the wave in hydrogen has been calculated and it is only about one-seventh that of the wave in air from the same disturbance, showing that the original disturbance is the more easily dissipated the lighter the medium.

It is not possible, however, to calculate the duration of the peak pressure in the wave or the pressure gradient in the rear of the wave, because these must depend upon the extent to which the wave is maintained. There is, however, one corollary to the theory of the shock wave, which is important in this connexion and which enables us to obtain a conception of the wave form under specified conditions. The pressure in the gas layers behind the wave front remains at the same peak value attained in the wave front for all those layers which continue to move at a speed U appropriate to the rate D at which the wave front is propagating. As soon as the velocity of any layer falls for any reason below U, so the pressure also falls below that at the wave front. This occurs for example when layers behind the wave front are affected by the rarefaction wave. The photographic records considered in conjunction with figure 6*b* enable us to obtain corroborative information on this point, if we assume as an approximation that the laws of continuous adiabatic compression are obeyed. It is possible then, by measurement of the distance travelled at any instant to determine the compression ratio and by means of the ordinary gas laws to calculate the pressure attained in the layers of air between the shock wave and the gases moving forward from the compression chamber, that is in the 'pressure wave'. With a diaphragm bursting at 134·7 lb./sq. in. absolute, and a compression chamber 61 cm. long, the uniform pressure in the pressure wave is calculated to be 44·9 lb./sq. in. absolute. According to Becker's equation the absolute pressure in the shock wave preceding the pressure wave is calculated from its speed (560 m./sec.) to be 46·8 lb./sq. in. This means that the pressure in the shock wave falls little, if at all (in view of the approximation used), in the distance defined as the pressure wave. With a still more intense disturbance, with a diaphragm bursting at 544·7 lb./sq. in. (abs.), the corresponding absolute pressures are 73·5 lb./sq. in. in the shock wave and 66·0 lb./sq. in. in the pressure wave. Again the agreement is good, for the error introduced by assuming continuous adiabatic compression in the pressure wave will increase in this sense as the intensity of the disturbance is increased. The distance over which the peak pressure in the shock wave will be maintained will depend on the length of the compression chamber, that is on the volume of maintaining gases. In the numerical examples discussed above, the figures suggest that the peak pressure in the shock wave is maintained by the gases in the pressure wave. The method can be extended to show that this state exists throughout the 67·7 cm. expansion chamber, when the compression chamber is 61 cm. long. Also that the pressure in the pressure wave persists through part of the column of the gases from the compression chamber, depending upon the position of the rarefaction wave in the compression chamber.

The pressure variations along the compression and expansion chambers can be represented graphically as in figure 21 *a*, *b*, which correspond with the instantaneous states depicted in respective diagrams in figure 6 *a*, *b*. The diagrams are drawn to be in keeping, qualitatively, with the results of the experimental work, except that the slope of the front of the shock wave W from the vertical is greatly exaggerated.

When the compression chamber is short, the whole column of compressed air will have begun to expand before the shock wave has traversed the expansion chamber. The sequence of events, showing the degeneration of the shock wave, is shown in figure 22 *a*–*h*. It is between stages *e* and *f* that the rate of advance of the shock wave begins to fall appreciably as a result of the decrease in the pressure ratio P_2/P_1. The sequence of events would be similar with longer compression and expansion chambers.

Calculation of speed of shock wave from bursting pressure. The comparison of calculated and experimental values for the velocity of the shock wave, D, and the velocity of the gas immediately behind it, U, may be taken as a verification of the theoretical relationships which have been advanced,

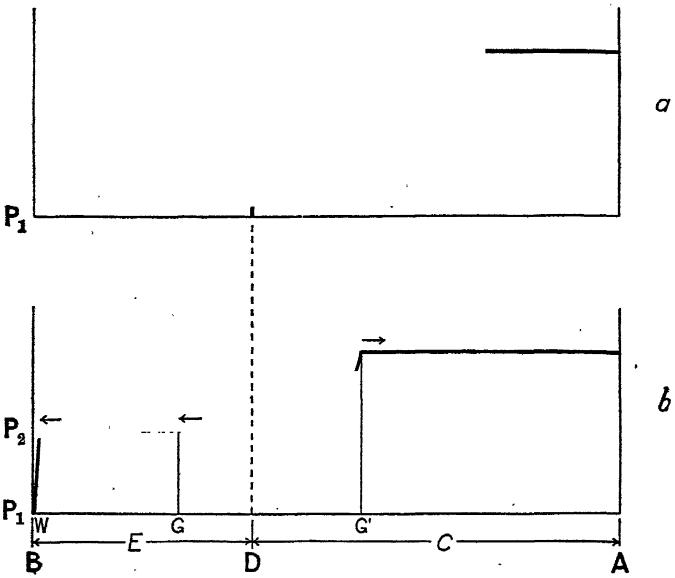


Figure 21. Diagrammatic representation of the pressure levels in the apparatus, (*a*) before the diaphram ruptures, (*b*) when the shock wave reaches end *B*.

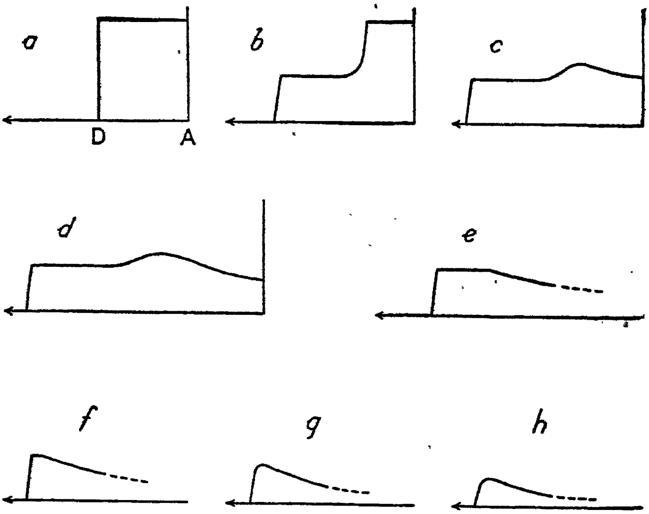


Figure 22. The decay of shock wave from a weak disturbance.

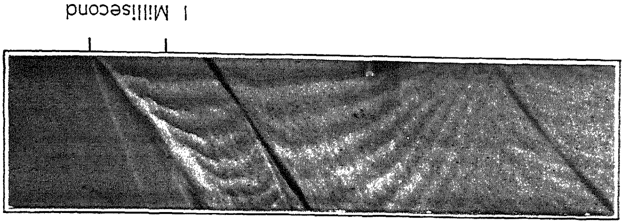


FIGURE 18

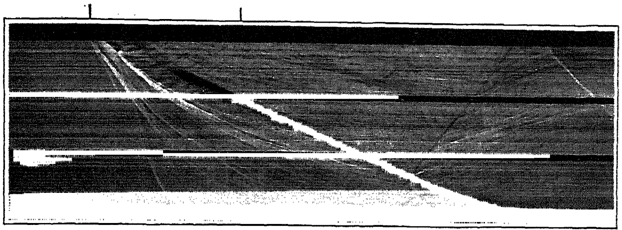


FIGURE 5

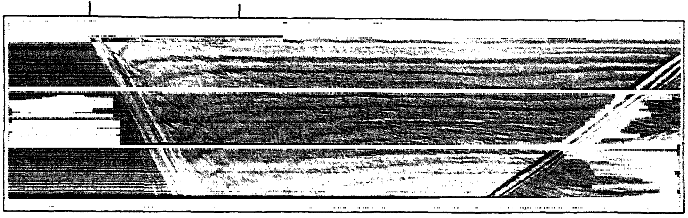


FIGURE 4

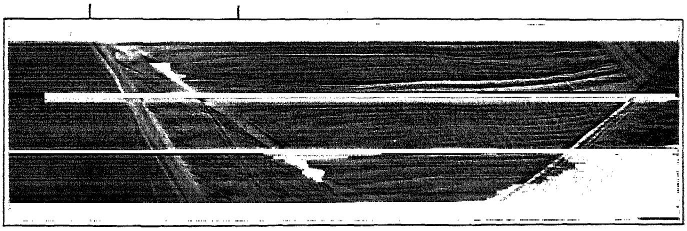


FIGURE 3

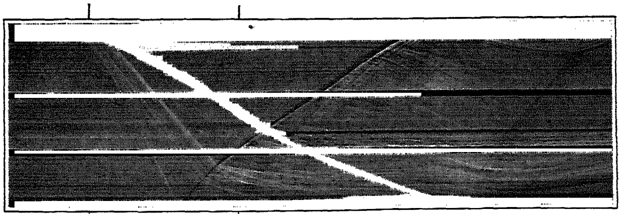


FIGURE 2

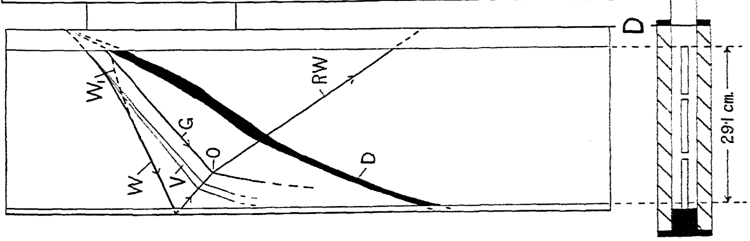


FIGURE 2a

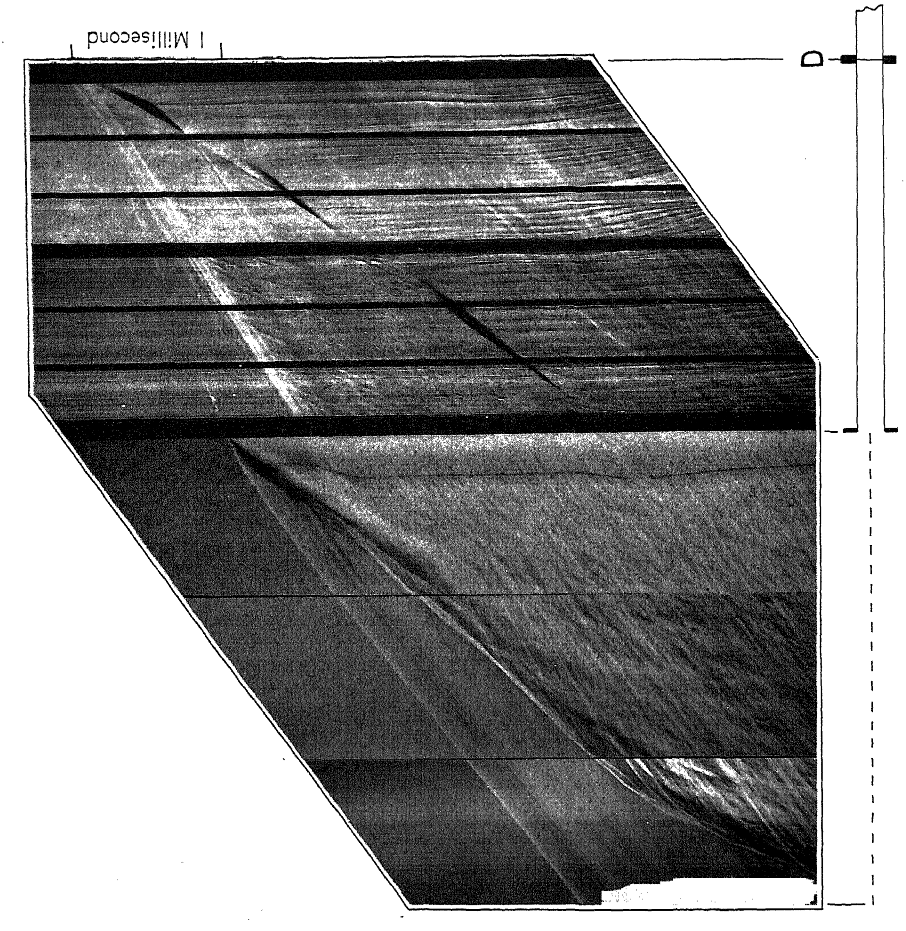


FIGURE 9

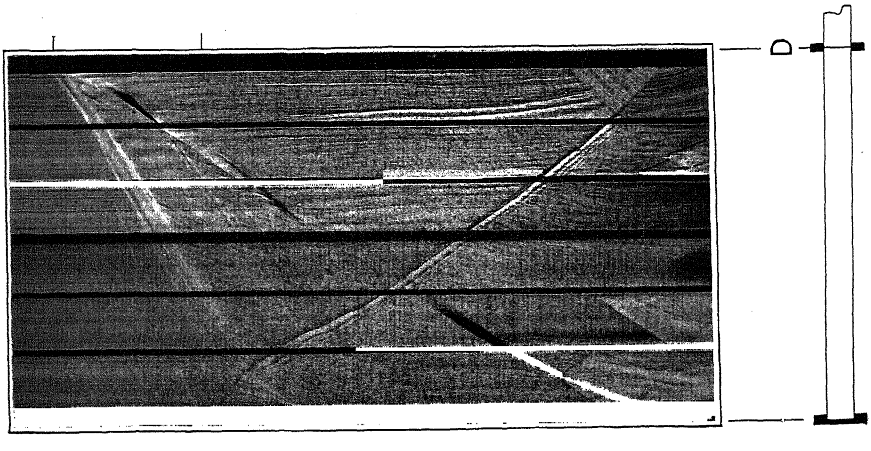


FIGURE 8

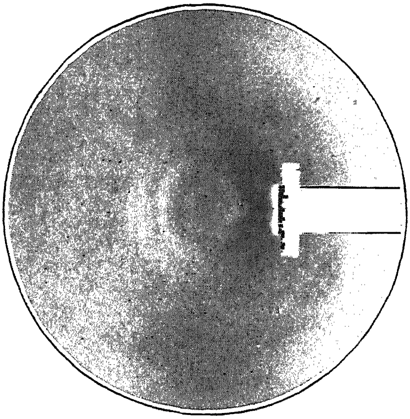


FIGURE 10

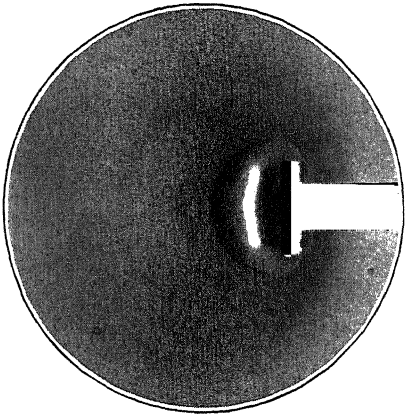


FIGURE 11

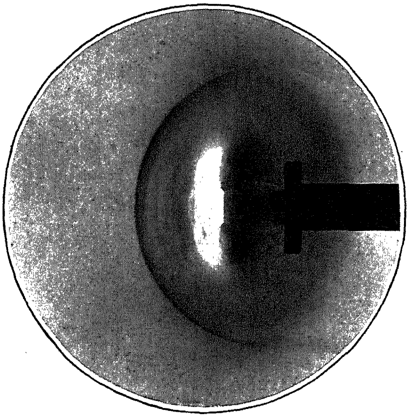


FIGURE 12

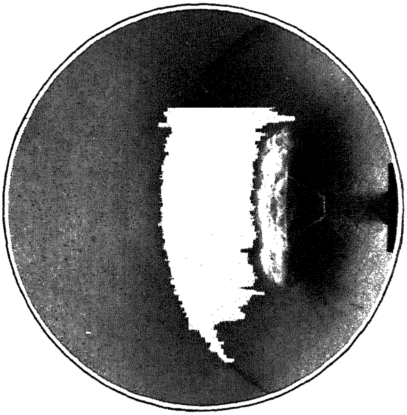


FIGURE 13

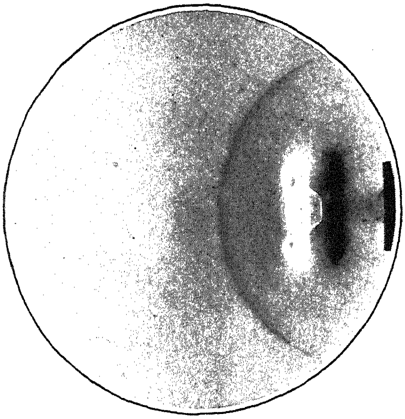


FIGURE 14

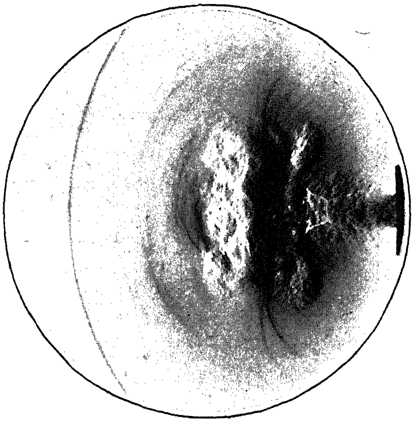


FIGURE 15

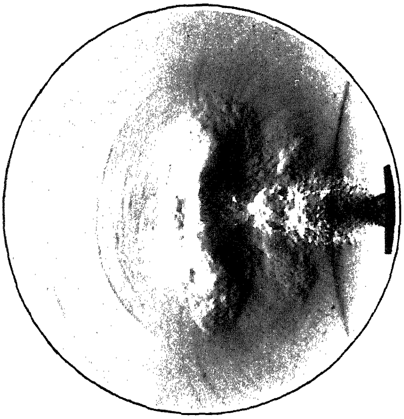


FIGURE 16

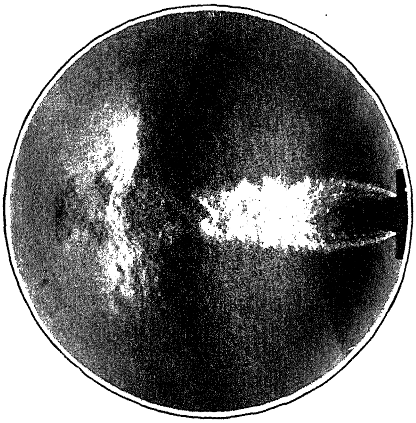


FIGURE 17

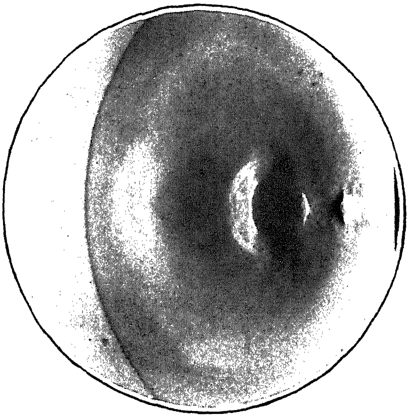


FIGURE 19

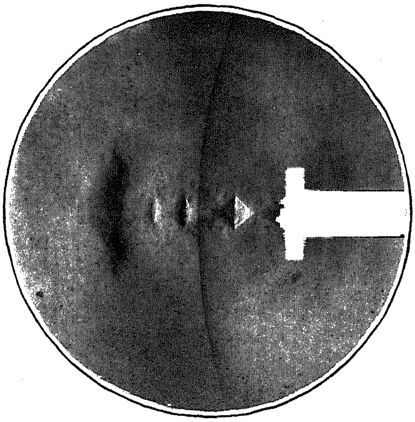


FIGURE 20

or, if the theory is regarded as established, as a confirmation of the accuracy of the experimental method. A further theoretical comparison which is of even greater value, suggested to us by Professor G. I. Taylor, is that between the pressure at which the diaphragm bursts and the speed of the stable shock wave established when the compression and expansion chambers are long. The treatment put forward by Professor Taylor is as follows. The symbols used have the same significance as in the earlier treatment; P_3 is the absolute pressure of the medium (air) in chamber C, T_1 its temperature. The speed of sound, E, in the compressed air is the same as that in the air in the expansion chamber.

The values of D and U are governed by the ratio $\pi = P_2/P_1$ and the relationship between D and P_3/P_1 is developed as follows. Let us consider the wave of rarefaction which moves into the compression chamber when the diaphragm ruptures. The plane which divides the air originally on opposite sides of the diaphragm moves at speed U (into the expansion chamber); the pressure on both sides of this plane is P_2 but the temperature T_2 on the shock wave side of the plane is above T_1, and the temperature T_3 on the rarefaction side is below T_1. In the rarefaction wave, the condition (1) is satisfied:

$$P_2 V_2^\gamma = P_3 V_3^\gamma, \tag{1}$$

where $V_3 = 1/\rho_3$ = specific volume of the undisturbed medium in chamber C. The wave is a progressive wave without shock surfaces and the movement U of the air left behind by the wave satisfies the Riemann condition for progressive waves of finite amplitude. This condition is given in (2):

$$U = \frac{2E}{\gamma - 1}\left\{1 - \left(\frac{P_2}{P_3}\right)^{(\gamma-1)/2\gamma}\right\}. \tag{2}$$

This relationship enables us to calculate the values of P_3/P_2 for various arbitrary values of U. The formulae given in the earlier portion of the paper give the values of P_2/P_1 for the same values of U and also give the related values of D. Finally, since $P_3/P_2 \times P_2/P_1 = P_3/P_1$ and $P_1 = 1$ atm., the theoretical values of D for various initial pressures in the compression chamber can be obtained. Calculations based on this treatment are given in the following table, initial conditions being taken as $T_1 = 273°$ K and $P_1 = 1$ atm.

On figure 7 are shown the calculated values of D, plotted against the bursting pressure of the diaphragm, and it will be seen that there is only a small divergence between theory and experiment. Some of this divergence arises from the fact that the experimental results were obtained at room temperature (10–15° C), the actual speeds on this account being 2–3 % higher than would be obtained at 0° C, the temperature used for all theoretical calculations in this paper.

U m./sec.	P_3/P_2	P_2/P_1	P_3 atm. abs.	D m./sec.
200	2·42	2·18	5·28	471
232	2·80	2·46	6·89	498
266	3·30	2·77	9·14	527
299	4·0	3·11	12·4	556
332	4·79	3·46	16·6	585
365	5·70	3·86	22·0	617
398	6·78	4·27	28·9	648
432	8·08	4·71	38·1	679
465	9·67	5·19	50·1	711
498	11·7	5·71	66·8	745

The writers thank the Safety in Mines Research Board for permission to publish this contribution which describes part of the research work on explosives done for the Board.

REFERENCES

Airy, G. B. 1849 *Phil. Mag.* **34**, 401–405.
Banerji, S. K. and Barave, R. V. 1931 *Phil. Mag.* **11**, 1057–1081.
Becker, R. 1917 *Z. Elektrochem.* **23**, 40–49.
Becker, R. 1921 *Z. Phys.* **4**, 393–409.
Becker, R. 1922 *Z. Phys.* **8**, 321–362.
Bonyun, M. E. 1935 *Trans. Amer. Inst. Chem. Engrs*, **31**, 1–17.
Chapman, D. L. 1899 *Phil. Mag.* **47**, 90–104.
Dixon, H. B. 1903 *Phil. Trans.* A, **200**, 315–353.
Duhem, P. 1909 *Z. Phys. Chem.* **69**, 169–186.
Earnshaw, S. 1859 *Proc. Roy. Soc.* A, **9**, 590–591.
Earnshaw, S. 1860 *Phil. Trans.* A, **150**, 133–148.
Emden 1899 *Ann. Phys., Lpz.*, **69**, 264–289, 426–453.
Gawthrop, D. B., Perrott, G. St J. and Shepherd, W. C. F. 1931 *J. Franklin Inst.* **211**, 67–86.
Hadamard 1903 *Propagation des Ondes.* Paris.
Hartmann, J. 1931 *Phil. Mag.* **11**, 926–948.
Hugoniot, H. 1887 *J. Ec. polyt. Paris*, **57**, 3–97; **58**, 1–125.
Jouguet, E. 1917 *Mécanique des Explosifs.* Paris.
Payman, W. 1928 *Proc. Roy. Soc.* A, **120**, 90–109.
Payman, W. and Robinson, H. 1926 *Pap. Safety Min. Res. Bd, Lond.*, No. 18.
Payman, W. and Shepherd, W. C. F. 1937 *Proc. Roy. Soc.* A, **158**, 348–367.
Payman, W. and Titman, H. 1935 *Proc. Roy. Soc.* A, **152**, 418–455.
Payman, W. and Woodhead, D. W. 1931 *Proc. Roy. Soc.* A, **132**, 200–213.
Poisson, M. 1808 *J. Ec. polyt. Paris*, **7**, 319–392.

Prandtl, L. 1904 *Phys. Z.* **5**, 599–601.
Prandtl, L. 1907 *Z. Phys.* **8**, 23–31.
Rankine, W. J. M. 1870 *Phil. Trans.* A, **160**, 277–288.
Rayleigh, Lord 1910 *Proc. Roy. Soc.* A, **84**, 247–284.
Riemann, B. 1860 *Nachr. Ges. Wiss. Göttingen*, **8**.
Schuster, A. 1893 *Phil. Trans.* A, **184**, 152–154.
Stanton, T. E. 1926 *Proc. Roy. Soc.* A, **111**, 306–339.
Stokes, G. G. 1848 *Phil. Mag.* **33**, 349–356.
Taylor, G. I. 1910 *Proc. Roy. Soc.* A, **84**, 371–377.
Templen 1905 *Enzykl. Math. Wiss.* **4**, 281–323.
Vieille, P. 1899 *Mem. Poudres et Salpêtres*, **10**, 177–260.

The diffraction of sound pulses
I. Diffraction by a semi-infinite plane

By F. G. Friedlander

(*Communicated by G. I. Taylor, F.R.S.—Received* 29 *February* 1940)

This paper contains the results of some calculations which show the changes undergone by a sound pulse when it is diffracted by an infinite screen or wall with a straight edge. The incident pulse is travelling in such a manner that its wave front is parallel to the plane of the wall and the motion is assumed to be two-dimensional. The calculations are carried out for a certain pulse in which the pressure rises instantaneously and then decays exponentially, and—in less detail—for several other types of incident pulse. The pressure changes in the geometrical shadow and near its boundary are investigated, as well as the pressure at points on the screen itself. A remarkable feature is the propagation of the initial pressure discontinuity along the boundary of the geometrical shadow as an instantaneous pressure discontinuity across this boundary. The problem could be treated by the application of Fourier transforms to Sommerfeld's well-known solution of the diffraction of simple harmonic waves by a straight edge, but the analysis utilized in this paper offers many advantages, particularly when the incident pulse starts with a discontinuous pressure rise. It is also shown that, although the two solutions of the problem treated in this paper (which are due to Sommerfeld and Lamb respectively) differ in form, one can be obtained from the other by a suitable transformation.

Introduction

The diffraction of a simple harmonic wave train by a straight-edged semi-infinite screen, and by an infinite wedge, was originally discussed by Sommerfeld (1895). In a later paper Sommerfeld (1901) extended his analysis to the diffraction of pulses and treated the case of a 'rectangular pulse' in detail. The analysis of Sommerfeld's first paper was simplified by Lamb (1906), who also dealt with the reflexion of simple harmonic waves by a convex parabolic cylinder, and by a paraboloid. Lamb then applied his method to the diffraction of a pulse by a semi-infinite screen. His result differs in form from that given by Sommerfeld, but it will be shown in this paper that it can be transformed in such a manner that Sommerfeld's formulae are obtained.

Although the cases of pulse diffraction and of 'wave' diffraction (the latter refers to infinite trains of simple harmonic waves) are analogous, they differ in several respects. The equation governing the propagation of pulses, i.e. the sound equation, is of the hyperbolic type; that governing wave diffraction is of the elliptic type. The functions satisfying these equations behave quite differently.* In wave diffraction, interference between 'elementary waves' produces certain alternating maxima and minima of intensity, which give rise to the diffraction bands well known in optics. Nothing of this nature must be expected in the case of a pulse: interference occurs because of the persistence and the periodicity of a wave train.

It is well known that every pulse can be thought of as made up by the superposition of simple harmonic waves; in other words, that it is possible to extend results about simple harmonic waves to the case of a pulse by applying a Fourier transform. When using this method during some preliminary calculations it was found that, if the integrals involved had to be evaluated numerically, the work tended to become very laborious and inaccurate. The formulae which were finally used were of a simpler type and much more suitable for numerical evaluation. They indicate very clearly the superiority of direct solutions of pulse diffraction problems over those obtained by Fourier transforms, particularly when the incident pulse has a well-defined wave front. It is usually difficult to obtain the general characteristics of the diffracted pulse from solutions of the Fourier transform type. For example, it requires considerable mathematical ingenuity to deduce the shape of the wave front in the shadow of an obstacle from such a solution, while it can be obtained from the direct solution by inspection. These difficulties are, on the whole, due to the fact that the Fourier transform provides the link between the solutions of partial differential equations of different types satisfied by different classes of functions; a point which has been mentioned above.

There are very few general methods available for obtaining the solutions of pulse diffraction problems. Similarly, approximate methods have only been little developed. Most of the approximate methods used in the treatment of wave diffraction apply only to limited frequency ranges and thus cannot be generalized by the application of a Fourier transform. Huygens's principle, in the form given to it by Kirchhoff,† can be applied to pulse diffraction.‡ But it seems unavoidable to ignore reflected waves in the application of this principle; this is equivalent to assuming the diffracting obstacles to be absorbing, an assumption which is quite legitimate in optics, but of doubtful validity in the theory of sound. As a material usually only absorbs sound whose wave-length (or some equivalent dimension) is a small fraction of the thickness of the material, Huygens's principle—together with the assumptions made in approximate calculations—must be applied with great caution.

---

* This is discussed, for example, by Hadamard in *Le problème de Cauchy* (Paris, 1932), pp. 37 *et seq.*

† See, for example, Lamb, *Hydrodynamics*, 6th ed. p. 501.

‡ I am indebted to Dr E. C. Bullard for pointing out this possibility.

The solutions of pulse diffraction problems which are known thus become important as an indication of the general character of pulse diffraction. It is hoped that the results given in this paper will serve, apart from their intrinsic interest, to give some general indications of this kind.

MATHEMATICAL SOLUTION

Choose the co-ordinate system so that the z-axis coincides with the edge of the wall, which occupies the positive part of the x-z-plane. The motion will then depend only on x, y, and the time t. It is convenient to introduce parabolic co-ordinates, defined by

$$x+iy = (\xi+i\eta)^2 \quad \text{or} \quad \left.\begin{aligned} \xi^2-\eta^2 &= x, \\ 2\xi\eta &= y. \end{aligned}\right\} \tag{1}$$

Thus the x-y-plane is transformed into the upper half of the ξ-η-plane, the positive half of the x-axis (which is the intersection of the wall with the x-y-plane) being transformed into the whole of the ξ-axis. The curves $\xi = \text{const.}$, $\eta = \text{const.}$ respectively are systems of confocal parabolas having the origin as common focus and the x-axis as common axis (see figure 1). The boundary condition expressing reflexion at the wall is

$$\frac{\partial \Phi}{\partial y} = 0 \quad (y=0, x \geqslant 0) \qquad \text{or} \qquad \frac{\partial \Phi}{\partial \eta} = 0 \quad (\eta = 0) \tag{2}$$

(see Lamb 1906), where Φ is the velocity potential. Thus in the transformed plane the whole of the ξ-axis appears as a reflector.

The velocity potential must satisfy the sound equation

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2}, \tag{3}$$

where c is the velocity of sound. It is assumed that the motion is sufficiently small for this equation to hold. The incident pulse can then be represented by the velocity potential

$$F(ct+y) \quad \text{or} \quad F(ct+2\xi\eta). \tag{4}$$

Lamb (1910) showed that the solution which can be made to satisfy the boundary conditions at the half-plane and at infinity is

$$\Phi = \tfrac{1}{2}F(ct+2\xi\eta)+\tfrac{1}{2}F(ct-2\xi\eta)+\int_0^{\xi+\eta} f(ct+2\xi\eta-\zeta^2)\,d\zeta + \int_0^{\xi-\eta} f(ct-2\xi\eta-\zeta^2)\,d\zeta \qquad (5)$$

(see Lamb 1910), where the function f, which appears under the integral sign, is unspecified. It is easily verified by differentiation that (5) satisfies the equation (3). It also fulfils the boundary condition (2). For we have

$$\frac{\partial\Phi}{\partial\eta} = \xi\{F'(ct+2\xi\eta)-F'(ct-2\xi\eta)\} + \int_0^{\xi+\eta} 2\xi f'(ct+2\xi\eta-\zeta^2)\,d\zeta - \int_0^{\xi-\eta} 2\xi f'(ct-2\xi\eta-\zeta^2)\,d\zeta,$$

which vanishes when $\eta = 0$. That this must be so is also apparent from the form of (5) which consists of two groups of terms, one of which is the 'image' of the other with respect to the ξ-axis.

To determine the function f, Lamb (1906) uses the additional condition that in the region of great negative x the motion must closely approximate to the incident wave, i.e. as

$$x \to -\infty, \quad \Phi \to F(ct+y). \qquad (6)$$

It follows from (1) that as

$$x \to -\infty, \quad \xi+\eta \to +\infty, \quad \xi-\eta \to -\infty,$$

hence (6) becomes, substituting from (5),

$$F(ct+2\xi\eta) = \tfrac{1}{2}F(ct+2\xi\eta)+\tfrac{1}{2}F(ct-2\xi\eta) + \int_0^{\infty} f(ct+2\xi\eta-\zeta^2)\,d\zeta - \int_0^{\infty} f(ct-2\xi\eta-\zeta^2)\,d\zeta.$$

For this to hold, f must satisfy the integral equation

$$\int_0^{\infty} f(z-\zeta^2)\,d\zeta = \tfrac{1}{2}F(z).^* \qquad (7)$$

---

* The equation should really be $\int_0^{\infty} f(z-\zeta^2)\,d\zeta = \tfrac{1}{2}F(z)+\text{const.}$, but it will be seen in the following (cf. (8)) that the constant has no effect on the result and can thus be ignored.

This can be transformed into Abel's integral equation, by putting

$$z-\zeta^2 = u, \quad d\zeta = -\frac{du}{2\sqrt{(z-u)}},$$

so that

$$\int_{-\infty}^{z} \frac{f(u)\,du}{\sqrt{(z-u)}} = F(z). \tag{7a}$$

Then, assuming that $\lim_{z=-\infty} F(z) = 0$, the solution of this integral equation is (Bocher 1909)

$$f(u) = \frac{1}{\pi}\int_{-\infty}^{u} \frac{F'(z)\,dz}{\sqrt{(u-z)}}.^* \tag{8}$$

Now consider the expression

$$\Phi_1 = \tfrac{1}{2}F(ct+2\xi\eta) + \int_0^{\xi+\eta} f(ct+2\xi\eta-\zeta^2)\,d\zeta. \tag{9}$$

This is part of (5). It is easily verified that it satisfies the sound equation (3). It can be written differently. Introducing polar co-ordinates

$$x+iy = re^{i\theta}, \tag{10}$$

it is seen that

$$\xi+\eta = +\sqrt{(r+y)}, \quad (0 \leqslant \theta \leqslant \tfrac{3}{2}\pi); \qquad \xi+\eta = -\sqrt{(r+y)}, \quad (\tfrac{3}{2}\pi \leqslant \theta \leqslant 2\pi);$$

hence

$$\Phi_1 = \tfrac{1}{2}F(ct+y) + \int_0^{\sqrt{(r+y)}} f(ct+y-\zeta^2)\,d\zeta, \quad (0 \leqslant \theta \leqslant \tfrac{3}{2}\pi)$$

$$= \tfrac{1}{2}F(ct+y) - \int_0^{\sqrt{(r+y)}} f(ct+y-\zeta^2)\,d\zeta, \quad (\tfrac{3}{2}\pi \leqslant \theta \leqslant 2\pi).$$

Taking account of (7) this can be put into the form

$$\Phi_1 = \int_{\sqrt{(r+y)}}^{\infty} f(ct+y-\zeta^2)\,d\zeta + [F(ct+y)], \tag{9a}$$

where the brackets indicate that $F(ct+y)$ must be omitted when $\tfrac{3}{2}\pi \leqslant \theta \leqslant 2\pi$. Substituting for f from (8),

$$\int_{\sqrt{(r+y)}}^{\infty} f(ct+y-\zeta^2)\,d\zeta = \frac{1}{\pi}\int_{\sqrt{(r+y)}}^{\infty} d\zeta \int_{-\infty}^{ct+y-\zeta^2} \frac{F'(u)\,du}{\sqrt{(ct+y-\zeta^2-u)}}.$$

This is a double integral over u and ζ, extended over the area to the right of $\zeta = \sqrt{(r+y)}$, and bounded by the parabola $u = ct+y-\zeta^2$. Inverting the order of integration and making the appropriate changes in the limits of integration,

---

* Lamb (1910) gave this solution in a footnote and utilized a different solution based on the application of a Fourier transform. (8) is more convenient in applications when f cannot be found explicitly.

$$\int_{\sqrt{(r+y)}}^{\infty} f(ct+y-\zeta^2)\,d\zeta = \frac{1}{\pi}\int_{-\infty}^{ct-r} F'(u)\,du \int_{\sqrt{(r+y)}}^{\sqrt{(ct-y+u)}} \frac{d\zeta}{\sqrt{(ct+y-u-\zeta^2)}}$$

$$= \frac{1}{\pi}\int_{-\infty}^{ct-r} F'(u)\left\{\tfrac{1}{2}\pi - \sin^{-1}\sqrt{\frac{r+y}{ct+y-u}}\right\} du.$$

Thus by a partial integration,

$$\int_{\sqrt{(r+y)}}^{\infty} f(ct+y-\zeta^2)\,d\zeta = \frac{\sqrt{(r+y)}}{2\pi}\int_{-\infty}^{ct-r} \frac{F(u)\,du}{(ct+y-u)\sqrt{(ct-r-u)}},$$

and hence

$$\Phi_1 = \frac{\sqrt{(r+y)}}{2\pi}\int_{-\infty}^{ct-r} \frac{F(u)\,du}{(ct+y-u)\sqrt{(ct-r-u)}} + [F(ct+y)], \tag{11}$$

which agrees with the result given by Sommerfeld (1901, p. 40).

FORMULAE FOR THE PRESSURE. APPROXIMATIONS

In the preceding section the solution of the problem of the diffraction of a pulse of arbitrary shape by a semi-infinite wall has been developed in terms of the velocity potential. In the following we shall be mainly concerned with the pressure, so that it is best to transcribe the solution obtained into terms of pressure. The pressure, p (i.e. the pressure in excess over that of the undisturbed atmosphere), is connected with the velocity potential by the relation

$$p = \rho\frac{\partial\Phi}{\partial t}. \tag{12}$$

Now it was found that

$$\Phi = \tfrac{1}{2}F(ct+y) + \tfrac{1}{2}F(ct-y) + \int_0^{\xi+\eta} f(ct+y-\zeta^2)\,d\zeta + \int_0^{\xi-\eta} f(ct-y-\zeta^2)\,d\zeta.$$

The function f is continuous, so that differentiation under the integral sign is permissible. Thus

$$p = \rho c\left\{\tfrac{1}{2}F'(ct+y) + \tfrac{1}{2}F'(ct-y) + \int_0^{\xi+\eta} f'(ct+y-\zeta^2)\,d\zeta + \int_0^{\xi-\eta} f'(ct-y-\zeta^2)\,d\zeta\right\}. \tag{13}$$

From now on we shall assume that the incident pulse has a definite wave front, so that we can put

$$F(z) = 0 \quad (z \leqslant 0). \tag{14}$$

Then the equations (7*a*) and (8) become

$$f(u) = \frac{1}{\pi}\int_0^u \frac{F'(z)\,dz}{\sqrt{(u-z)}}; \quad F(z) = \int_0^z \frac{f(u)\,du}{\sqrt{(z-u)}}. \tag{15}$$

To effect the differentiation of f, a partial integration is performed first. Thus it is found that

$$f(u) = \frac{1}{\pi}\int_0^u \frac{F'(z)\,dz}{\sqrt{(u-z)}} = \frac{1}{\pi}\left\{2\sqrt{u}\,F'(0) + 2\int_0^u \sqrt{(u-z)}\,F''(z)\,dz\right\},$$

hence

$$f'(u) = \frac{F'(0)}{\pi\sqrt{u}} + \frac{1}{\pi}\int_0^u \frac{F''(z)\,dz}{\sqrt{(u-z)}}, \tag{16}$$

as we cannot suppose that $F'(0) = 0$. Introducing a new auxiliary function ϕ and the pressure of the incident pulse p_0, by putting

$$\phi(u) = \rho c f'(u), \quad p_0(z) = \rho c F'(z), \tag{17}$$

the equations (13) and (16) can be written

$$\phi(u) = \frac{p_0(0)}{\pi\sqrt{u}} + \frac{1}{\pi}\int_0^u \frac{p_0'(z)\,dz}{\sqrt{(u-z)}}, \tag{16a}$$

$$\begin{aligned} p = {} & \tfrac{1}{2}p_0(ct+y) + \tfrac{1}{2}p_0(ct-y) \\ & + \int_0^{\xi+\eta} \phi(ct+y-\zeta^2)\,d\zeta + \int_0^{\xi-\eta} \phi(ct-y-\zeta^2)\,d\zeta. \end{aligned} \tag{13a}$$

The converse of (16*a*) is obtained by differentiating the second equation in (15) and introducing ϕ and p_0. Since $f(0) = 0$ it is found that

$$p_0(z) = \int_0^z \frac{\phi(u)\,du}{\sqrt{(z-u)}}. \tag{18}$$

Remembering that the transformation (1) transforms the whole of the x-y-plane into the upper half of the ξ-η-plane, and that therefore always $\eta > 0$, it is easily deduced that

$$\left.\begin{array}{llll} \text{(Region 1)} & \xi+\eta = +\sqrt{(r+y)}, & \xi-\eta = +\sqrt{(r-y)}, & (x>0, y>0; 0<\theta<\tfrac{1}{2}\pi), \\ \text{(Region 2)} & \xi+\eta = +\sqrt{(r+y)}, & \xi-\eta = -\sqrt{(r-y)}, & (x<0, \tfrac{1}{2}\pi<\theta<\tfrac{3}{2}\pi), \\ \text{(Region 3)} & \xi+\eta = -\sqrt{(r+y)}, & \xi-\eta = -\sqrt{(r-y)}, & (x>0, y<0, \tfrac{3}{2}\pi<\theta<2\pi). \end{array}\right\} \tag{19}$$

Hence making use of the equation (18), it is seen that the formula for the pressure, (13*a*), can be written as follows:

$$p_1 = p_0(ct+y)+p_0(ct-y)-\int_{\sqrt{(r+y)}}^{\sqrt{(ct+y)}} \phi(ct+y-\zeta^2)\,d\zeta-\int_{\sqrt{(r-y)}}^{\sqrt{(ct-y)}} \phi(ct-y-\zeta^2)\,d\zeta,$$

$$p_2 = p_0(ct+y)-\int_{\sqrt{(r+y)}}^{\sqrt{(ct+y)}} \phi(ct+y-\zeta^2)\,d\zeta+\int_{\sqrt{(r-y)}}^{\sqrt{(ct-y)}} \phi(ct-y-\zeta^2)\,d\zeta,$$

$$p_3 = \int_{\sqrt{(r+y)}}^{\sqrt{(ct+y)}} \phi(ct+y-\zeta^2)\,d\zeta+\int_{\sqrt{(r-y)}}^{\sqrt{(ct-y)}} \phi(ct-y-\zeta^2)\,d\zeta,$$

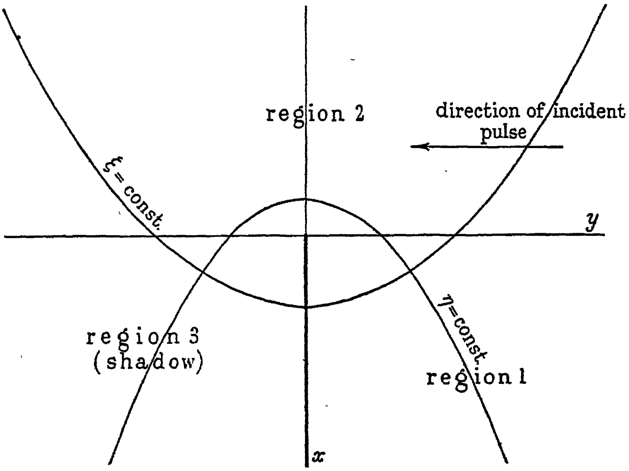


FIGURE 1

where the indices 1, 2, 3 indicate that these formulae apply in the regions labelled 1, 2, 3 in (19) (see also figure 1). Now the integrals appearing in these formulae can be transformed by the substitutions

$$ct+y-\zeta^2 = u, \quad \int_{\sqrt{(r+y)}}^{\sqrt{(ct+y)}} \phi(ct+y-\zeta^2)\,d\zeta = \frac{1}{2}\int_0^{ct-r} \frac{\phi(u)\,du}{\sqrt{(ct+y-u)}},$$

$$ct-y-\zeta^2 = u, \quad \int_{\sqrt{(r-y)}}^{\sqrt{(ct-y)}} \phi(ct-y-\zeta^2)\,d\zeta = \frac{1}{2}\int_0^{ct-r} \frac{\phi(u)\,du}{\sqrt{(ct-y-u)}}.$$

Then, introducing the function

$$P(X,T) = \int_0^T \frac{\phi(u)\,du}{\sqrt{(T+X-u)}}, \tag{20}$$

the following expressions for the pressure are obtained:

$$\left.\begin{aligned} p_1 &= p_0(ct+y)+p_0(ct-y)-\tfrac{1}{2}\{P(X_1,T)+P(X_2,T)\},\\ p_2 &= p_0(ct+y)-\tfrac{1}{2}\{P(X_1,T)-P(X_2,T)\},\\ p_3 &= \tfrac{1}{2}\{P(X_1,T)+P(X_2,T)\},\end{aligned}\right\} \tag{21}$$

where $$X_1 = r+y, \quad X_2 = r-y, \quad T = ct-r. \tag{21a}$$

It is possible to express the function P directly in terms of p_0. For, using (16a),

$$\begin{aligned} P(X,T) &= \frac{2p_0(0)}{\pi}\tan^{-1}\sqrt{\frac{T}{X}}+\frac{1}{\pi}\int_0^T p_0'(z)\,dz\int_z^T \frac{du}{\sqrt{[(u-z)(T+X-u)]}}\\ &= \frac{2p_0(0)}{\pi}\tan^{-1}\sqrt{\frac{T}{X}}+\frac{2}{\pi}\int_0^T p_0'(z)\tan^{-1}\sqrt{\frac{T-z}{X}}\,dz,\end{aligned}$$

hence $$P(X,T) = \frac{\sqrt{X}}{\pi}\int_0^T \frac{p_0(z)\,dz}{(T+X-z)\sqrt{(T-z)}}, \tag{20a}$$

by a partial integration.

Thus the pressure is given by (21) in terms of the incident pressure pulse p_0 and of the function P. It is useful to have approximate expressions for this function when X is either very large or very small. These can be obtained as follows:

(1) *X is very large.* The variation of $\sqrt{(T+X-u)}$ in the interval $(0,T)$ is small and so approximately

$$P(X,T) = \frac{1}{\sqrt{X}}\int_0^T \phi(u)\,du = \frac{1}{\sqrt{X}}\Psi(T). \tag{22}$$

It follows from (17) and (15) that

$$\Psi(T) = \frac{1}{\pi}\int_0^T \frac{p_0(z)\,dz}{\sqrt{(T-z)}}. \tag{23}$$

(2) *X is small.* We can write

$$P(X,T) = \int_0^{T+X}\frac{\phi(u)\,du}{\sqrt{(T+X-u)}}-\int_T^{T+X}\frac{\phi(u)\,du}{\sqrt{(T+X-u)}} = p_0(T+X)-I(X,T).$$

Then, by a partial integration,

$$I(X,T) = \int_T^{T+X}\frac{\phi(u)\,du}{\sqrt{(T+X-u)}} = 2\phi(T)\sqrt{X}+2\int_T^{T+X}\phi'(u)\sqrt{(T+X-u)}\,du.$$

An upper limit for the integral appearing after the partial integration has been performed is clearly given by

$$\left|2\int_{T}^{T+X}\phi'(u)\sqrt{(T+X-u)}\,du\right| \leqq 2X^{\frac{3}{2}}\,|\,\phi'_{\mathrm{max.}}(T,X)\,|,$$

where $|\,\phi'_{\mathrm{max.}}(T,X)\,|$ is the greatest value of $|\,\phi'(T)\,|$ in the interval $(T, T+X)$; if X and $|\,\phi'_{\mathrm{max.}}\,|$ are sufficiently small, the integral can be neglected and thus

$$P(X,T) = p_0(T+X) - 2\sqrt{X}\,\phi(T) \text{ approximately.} \tag{24}$$

It will be seen in the following that these approximations are very useful when calculating the pressure distribution near the boundary of the 'geometrical shadow' of the wall.

It follows from approximation (22) that when X is sufficiently large, $P(X,T)$ can be neglected. It is thus possible to discuss the behaviour of the diffracted pulse far away from the edge of the wall. In terms of polar co-ordinates

$$X_1 = r(1+\sin\theta), \quad X_2 = r(1-\sin\theta).$$

Now suppose that θ varies while r is kept fixed, r being very large. Assuming that P can be neglected if $X > X^*$ it appears that the terms in P in the equations (21) can be neglected everywhere except for a certain region containing the y-axis, whose boundary will be approximately given by the parabolas

$$r(1+\sin\phi) = X^*, \quad r(1-\sin\phi) = X^*. \tag{25}$$

Outside this region

$$\left.\begin{aligned} p_1 &= p_0(ct+y)+p_0(ct-y),\\ p_2 &= p_0(ct+y),\\ p_3 &= 0. \end{aligned}\right\} \tag{26}$$

The pressure varies continuously in the area bounded by (25) and we find that on the boundary of the shadow very far from the origin,

$$\left.\begin{aligned} &p \to \tfrac{1}{2}p_0(ct+y), && (x=0,\ y\to-\infty)\\ &p \to \tfrac{1}{2}p_0(ct+y)+\tfrac{1}{2}p_0(ct-y), && (x=0,\ y\to+\infty) \end{aligned}\right\} \tag{26a}$$

since

$$P(0,T) = p_0(T).$$

CALCULATIONS: THE PULSE $p_0 = (1-z)\,e^{-z}$

The theoretical results developed in the preceding sections have been applied to the detailed discussion of the diffraction of a pulse given by

$$\left.\begin{aligned} p_0(ct+y) &= \left(1-\frac{ct+y}{\lambda}\right)\exp\left[-\frac{ct+y}{\lambda}\right], && (ct+y>0) \\ &= 0 && (ct+y<0) \end{aligned}\right\} \tag{27}$$

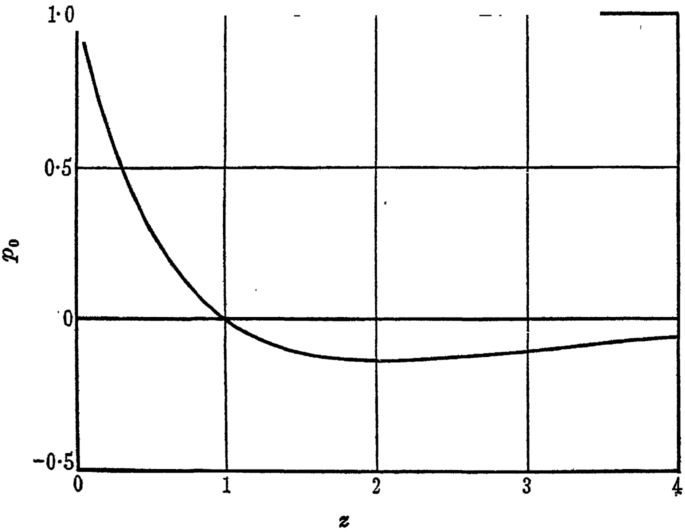


FIGURE 2. The pulse $p_0(z) = (1-z)\,e^{-z}$.

(see figure 2). λ is here a certain parameter characterizing the pulse which may be called the 'pulse thickness'. It is more convenient to use non-dimensional variables. We chose the pulse thickness as the unit of length, so that from now on x and y stand for x/λ, y/λ, i.e. they are the distances measured in terms of λ. The time t is replaced by the variable z which is connected with it by the relation

$$z = \frac{ct}{\lambda}. \tag{28}$$

(27) can then be rewritten as

$$\left.\begin{aligned} p_0(z+y) &= [1-(z+y)]\,e^{-(z+y)}, && (z+y>0) \\ &= 0. && (z+y<0) \end{aligned}\right\} \tag{27a}$$

The first step in the evaluation of the result is the calculation of the auxiliary functions Ψ, Ψ', Ψ''. It is found from (23) that

$$\Psi = \frac{1}{\pi}\int_0^T \frac{(1-z)\,e^{-z}}{\sqrt{(T-z)}}\,dz = \frac{2e^{-T}}{\pi}\int_0^{\sqrt{T}} (1-T+u^2)\,e^{u^2}\,du,$$

where the second form is easily obtained from the first by a change of variable. By a partial integration it is found that

$$\int_0^{\sqrt{T}} 2u^2 e^{u^2}\,du = -\int_0^{\sqrt{T}} e^{u^2}\,du + e^T\sqrt{T}.$$

Hence the following three expressions are obtained:

$$\Psi(T) = \frac{e^{-T}}{\pi}(1-2T)\int_0^{\sqrt{T}} e^{u^2}\,du + \frac{\sqrt{T}}{\pi}, \tag{29a}$$

$$\Psi'(T) = \frac{e^{-T}}{\pi}(2T-3)\int_0^{\sqrt{T}} e^{u^2}\,du + \frac{1}{\pi}\left(\frac{1}{\sqrt{T}} - \sqrt{T}\right), \tag{29b}$$

$$\Psi''(T) = \frac{e^{-T}}{\pi}(5-2T)\int_0^{\sqrt{T}} e^{u^2}\,du + \frac{1}{\pi}\left(\sqrt{T} - \frac{2}{\sqrt{T}} - \frac{1}{2T^{\frac{3}{2}}}\right). \tag{29c}$$

TABLE 1. THE FUNCTION $f(z) = (1-z)\,e^{-z}$

z	$f(z)$	z	$f(z)$	z	$f(z)$
0·00	1·000	1·25	−0·072	3·00	−0·100
0·25	0·583	1·50	−0·111	3·50	−0·075
0·50	0·303	1·75	−0·130	4·00	−0·057
0·75	0·115	2·00	−0·135	4·50	−0·039
1·00	0·000	2·50	−0·123	5·00	−0·027

TABLE 2. THE FUNCTIONS $\Psi(T)$, $\Psi'(T)$, $\Psi''(T)$ (cf. formulae (29*a*)–(29*c*))

T	$\Psi(T)$	$\Psi'(T)$	$\Psi''(T)$
0·0	0·000	∞	$-\infty$
0·2	0·217	0·246	−1·492
0·4	0·232	−0·038	−0·777
0·6	0·212	−0·146	−0·263
0·8	0·182	−0·170	−0·066
1·0	0·147	−0·171	0·036
1·5	0·070	−0·130	0·103
2·0	0·018	−0·081	0·088
2·5	0·000	−0·051	0·062
3·0	−0·028	−0·020	0·037
3·5	−0·034	−0·006	0·022
4·0	−0·035	0·002	0·011
4·5	−0·028	0·002	0·007
5·0	−0·026	0·005	0·002

TABLE 3. DIFFRACTED PRESSURE-TIME CURVES FOR THE INCIDENT PULSE

$$p_0 = (1-z)\,e^{-z}$$

NEAR THE BOUNDARY OF THE SHADOW, AT $y = -10$

(*a*) *outside the shadow*

τ	$x=-1\cdot0$	$x=-0\cdot8$	$x=-0\cdot6$	$x=-0\cdot4$	$x=-0\cdot2$	$x=0\cdot0$
0·0	1·000	1·000	1·000	1·000	1·000	0·500
0·2	0·443	0·444	0·416	0·402	0·388	0·360
0·4	0·236	0·241	0·230	0·221	0·193	0·232
0·6	0·107	0·110	0·114	0·122	0·128	0·146
0·8	0·032	0·039	0·044	0·052	0·056	0·069
1·0	−0·023	−0·009	−0·006	0·004	0·006	0·016
1·5	−0·080	−0·072	−0·063	−0·061	−0·054	−0·048
2·0	−0·088	−0·080	−0·072	−0·073	−0·070	−0·063
2·5	−0·071	−0·070	−0·068	−0·068	−0·068	−0·060
3·0	−0·054	−0·055	−0·056	−0·057	−0·054	−0·051

(*b*) *inside the shadow*

T	$x=0\cdot2$	$x=0\cdot4$	$x=0\cdot6$	$x=0\cdot8$	$x=1\cdot0$
0·0	0·000	0·000	0·000	0·000	0·000
0·2	0·337	0·321	0·306	0·289	0·275
0·4	0·228	0·226	0·225	0·218	0·210
0·6	0·139	0·143	0·146	0·148	0·148
0·8	0·072	0·079	0·085	0·089	0·092
1·0	0·024	0·030	0·036	0·042	0·046
1·5	−0·042	−0·037	−0·032	−0·026	−0·022
2·0	−0·058	−0·058	−0·055	−0·051	−0·046
2·5	−0·059	−0·057	−0·054	−0·052	−0·049
3·0	−0·052	−0·051	−0·050	−0·049	−0·047

TABLE 4. PRESSURE-TIME CURVES AT VARIOUS POINTS IN THE SHADOW

(incident pulse, $p_0(z) = (1-z)\,e^{-z}$)

T	$x=0\cdot75$ $y=0\cdot00$	$x=1\cdot50$ $y=0\cdot00$	$x=1\cdot50$ $y=-1\cdot50$	$x=1\cdot50$ $y=-3\cdot00$
0·00	0·000	0·000	0·000	0·000
0·25	0·242	0·179	0·180	0·205
0·50	0·212	0·210	0·176	0·175
0·75	0·149	0·167	0·138	0·133
1·00	0·098	0·098	0·091	0·084
1·50	0·012	0·026	0·021	0·013
2·00	−0·034	−0·004	−0·017	−0·021
2·50	−0·046	−0·026	−0·030	−0·040
3·00	−0·054	−0·036	−0·031	−0·038

TABLE 5. DIFFERENCE OF THE PRESSURES ON THE TWO SIDES OF THE WALL AT TWO POINTS

(incident pulse, $p_0(z) = (1-z)\,e^{-z}$)

z	$x = 0·75$	$x = 1·50$
0·00	2·000	2·000
0·25	1·164	1·164
0·50	0·606	0·606
0·75	0·229	0·229
1·00	−0·484	0·000
1·50	−0·621	−0·223
2·00	−0·378	−0·710
2·50	−0·206	−0·442
3·00	−0·119	−0·251
3·50	−0·046	−0·142
4·00	−0·001	−0·063

TABLE 6. DIFFRACTED PRESSURE-TIME CURVES FOR THE INCIDENT PULSE

$p_0(z) = 0 \quad (z \leqslant 0)$; $\qquad p_0(z) = 2z \quad (0 \leqslant z \leqslant \tfrac{1}{2})$;
$p_0(z) = 2(1-z) \quad (\tfrac{1}{2} \leqslant z \leqslant 1)$; $\qquad p_0(z) = 0 \quad (z \geqslant 1)$

NEAR THE BOUNDARY OF THE SHADOW, AT $y = -10$

(*a*) *outside the shadow*				(*b*) *inside the shadow*		
τ	$x = -1·0$	$x = -0·5$	$x = 0·0$	T	$x = 0·5$	$x = 1·0$
0·00	0·000	0·000	0·000	0·00	0·000	0·000
0·25	0·490	0·363	0·262	0·25	0·203	0·162
0·50	0·642	0·602	0·534	0·50	0·445	0·376
0·75	0·214	0·262	0·287	0·75	0·296	0·288
1·00	−0·058	−0·010	0·027	1·00	0·072	0·105
2·00	−0·005	0·016	0·024	2·00	0·034	0·042
3·00	0·015	0·023	0·020	3·00	0·025	0·029

Tables for

$$\int_0^Z e^{u^2}\,du$$

are available as far as $Z = 2$ (Jahnke-Emde 1933). These allow the calculation of Ψ, Ψ', Ψ'' as far as $T = 4$. For larger values of T, the integral was found by numerical integration. As great accuracy is not required over this range, the values of the functions (29) being small, this integration must be considered as a somewhat rough approximation to the exact value of the integral. The graphs of Ψ, Ψ', Ψ'' are shown in figure 3.

Figure 4 summarizes results relating to the behaviour of the pressure close to the edge of the geometrical shadow of the wall. A number of

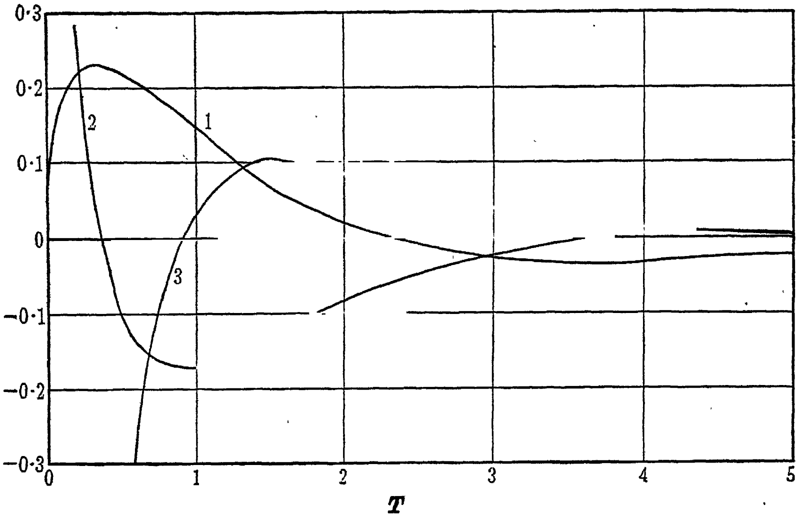


FIGURE 3. The functions Ψ, Ψ', Ψ'' for the incident pulse $p_0(z) = (1-z)\,e^{-z}$. The ordinate in curve 1 shows $\Psi(T)$, curve 2 $\Psi'(T)$, curve 3 $\Psi''(T)$.

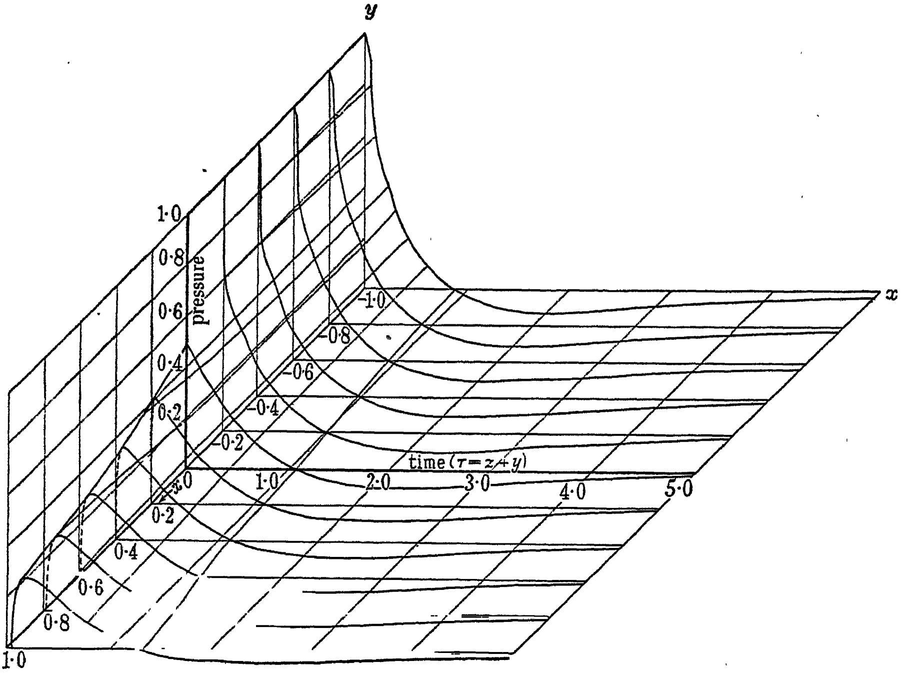


FIGURE 4. Pressure as a function of x and time at $y = -10$. The boundary of the geometrical shadow is $x = 0_x$.

pressure-time curves have been computed at points where $y = -10$, and x ranges from $+1$ to -1. These curves are superposed in the figure in such a way that they form the oblique projection of a surface whose perpendicular height represents the pressure. In the figure, the vertical ordinates represent the pressure; the horizontal ones the time; and the third co-ordinate measured parallel to the third (inclined) axis the distance from the edge of the shadow.

A remarkable feature, which will occur in all cases where a pulse with a discontinuous initial pressure rise is diffracted, is the initial pressure discontinuity across the boundary of the shadow which can be seen in the figure. The diffracted pulse consists of two components; one is a pulse with the cylindrical wave front $z = r$; the other one is, in the region under consideration, the incident pulse, which is propagated outside the shadow. Thus at any point outside the shadow the full and undistorted incident pulse will be experienced until the diffracted wave, which starts at the origin, arrives. The wave fronts touch at the boundary of the shadow; across this boundary the pressure is in general continuous, except initially. In fact, however close to the boundary a point may be, there will still be a small interval of time during which the incident pulse is experienced while the 'cylindrical' pulse has not yet arrived. If the incident pressure pulse starts discontinuously, then this discontinuity is thus propagated along the boundary of the shadow. If the pressure of the incident pulse rises continuously from the start, no discontinuity will occur at all (for an example cf. figure 9). It is evident from the nature of this phenomenon that it would not easily be detected if the solution employed were of the Fourier transform type.

Figure 4 shows clearly the attenuating effect of the wall: the maximum pressure reached falls off very quickly to about a third of the maximum pressure of the incident pulse.

The calculations were carried out by using the approximations (22) and (24) developed in the preceding section. These cannot be used to calculate the pressure immediately after the onset of the wave, however, as $\Psi'(T)$ and $\Psi''(T)$ are infinite when $T = 0$. For this purpose different approximations were obtained by replacing $(1-z)e^{-z}$ by $1-2z$. An application of $(20a)$ yields the result

$$P^*(X, T) = \frac{2}{\pi}\sqrt{(XT)} + \frac{2}{\pi}(1 - 2T - X)\tan^{-1}\sqrt{\frac{T}{X}}. \tag{30}$$

The asterisk indicates that this is only an approximation. (30) was only used for $0 \leqslant T \leqslant 0{\cdot}2$. For the range $T \geqslant 0{\cdot}4$ the approximations (22) and (24) were applied. Let

$$\tau = z+y$$

denote the time reckoned from the onset of the pulse at any point outside the shadow; then $\tau = T+X_1$. Hence if (22) and (24) are introduced in the formulae for the pressure, (21), it is found that the approximate expression for the pressure is:

$$\left.\begin{aligned}&\text{in the shadow}\\ &\qquad p^* = \frac{1}{2}\left\{p_0(\tau)+\frac{1}{\sqrt{X_2}}\Psi(\tau-X_1)-2\sqrt{X_1}\,\Psi'(\tau-X_1)\right\},\\ &\text{outside the shadow}\\ &\qquad p^* = \frac{1}{2}\left\{p_0(\tau)+\frac{1}{\sqrt{X_2}}\Psi(\tau-X_1)+2\sqrt{X_1}\,\Psi'(\tau-X_1)\right\}.\end{aligned}\right\} \tag{31}$$

Two ways are now open of calculating the pressure without approximation (except that inherent in numerical integration). Either the formula (20) for P can be used in conjunction with the table of $\Psi' = \phi$, or formula (20a) may be used. The second method is more direct, but the first one has the advantage of furnishing the values of Ψ, Ψ', Ψ'' which are needed in the approximations. If it is used, the contribution due to the term $\frac{1}{\pi}\frac{1}{\sqrt{u}}$ in (29b) must be evaluated separately. It is

$$\frac{2}{\pi}\tan^{-1}\sqrt{\frac{T}{X}}. \tag{32}$$

The contribution of the term $\pi\sqrt{u}$ to the result can also be evaluated explicitly, but it is more conveniently included in the numerical integration.[†]

To obtain a more general survey of the shadowing effect of the wall, lines of equal maximum pressure in the shadow were plotted. They are shown in figure 5. At any point, e.g. on the line labelled 0·2, the maximum pressure will be exactly 0·2. The 0·5 line coincides with the y-axis (boundary of the shadow) from $y = -\infty$ until some point near the origin (whose position was not calculated: it is obviously quite close to the origin), where it departs from the y-axis; inside the small space bounded by the y-axis, the x-axis, and this part of the 0·5 line the maximum pressure rises from 0·5 to its value at the origin, which is 1·0. To plot the curves shown in figure 5, the maximum pressure at a network of points in the shadow was calculated by computing the pressure for a sufficiently long time interval at each point; the curves were then obtained by interpolation. The network consisted of the points

---

[†] This is the method which was actually used in nearly all the calculations described in this paper. The reason is somewhat accidental; formulae (21), etc., were developed from the analysis in Lamb's paper (1910), as the solution given there was found unsuitable for numerical work. The bulk of the calculations had been completed when the connexion with Sommerfeld's work was realized and the formulae (20a) and (11) obtained. It will be found that if (20a) is used in the calculations described in this section, the integral $\int_0^Z e^{u^2}\,du$ will appear, so that Ψ, Ψ', Ψ'' could be computed.

$$\begin{aligned} &y = -20, x = 1; \quad y = -10, x = 1; \quad y = -5, x = 1; \\ &y = -20, x = 2; \quad y = -10, x = 2; \quad y = -5, x = 2; \\ &y = -20, x = 3; \quad y = -10, x = 3; \quad y = -5, x = 3. \end{aligned}$$

These lines of equal maximum pressure all tend asymptotically to parabolas of the form $r+y =$ const. but the range of y in figure 5 is not large enough to permit a verification of this.

It appears that the maximum pressure reached falls off very quickly to about a third of its value in the pulse, but decreases considerably less rapidly afterwards.

Pressure-time curves at some points in the shadow are shown in figure 6. It will be noticed that the decrease in maximum pressure is accompanied by a lengthening of the time interval during which the pressure is positive (it must be remembered that the 'pressure' is actually the excess pressure over that of the undisturbed atmosphere). If any pressure-time curves are compared it will be found that the curve with the lower maximum intersects the time-axis farther away from the origin as the other curve.

The curves in figure 6 may be compared with those in figure 7, showing the pressure variations at points, also in the shadow, at the back of the wall itself. The general shapes of all these curves are very similar.

The pressure on the front of the wall, p_1, is connected with that on the back of the wall, p_3, by the relation

$$p_1 + p_3 = 2p_0$$

which follows from (21). Hence the difference of the pressures on the two sides of the wall is at any point

$$p_1 - p_3 = 2(p_0 - p_3).$$

This pressure difference, for the points $x = 0{\cdot}75$ and $x = 1{\cdot}50$, is shown in figure 8. Due to the reflexion, the pressure is equal to twice the pressure

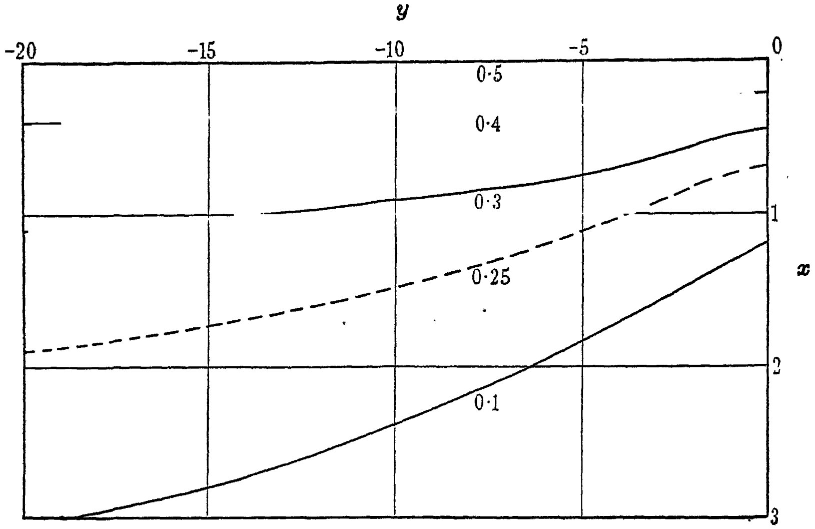


FIGURE 5. Lines of equal maximum pressure in the shadow. Note: the 0·5 line departs from the y-axis at a point near the origin.

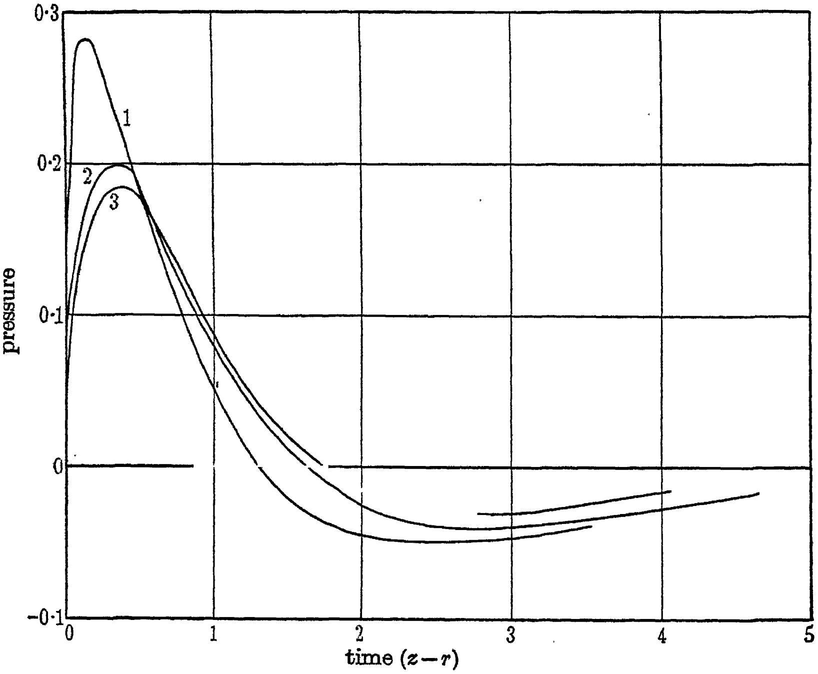


FIGURE 6. Pressure-time curves at some points in the shadow. Curve 1, $x = 1\cdot0$, $y = -10$; curve 2, $x = 1\cdot5$, $y = -3\cdot0$; curve 3, $x = 1\cdot5$, $y = -1\cdot5$.

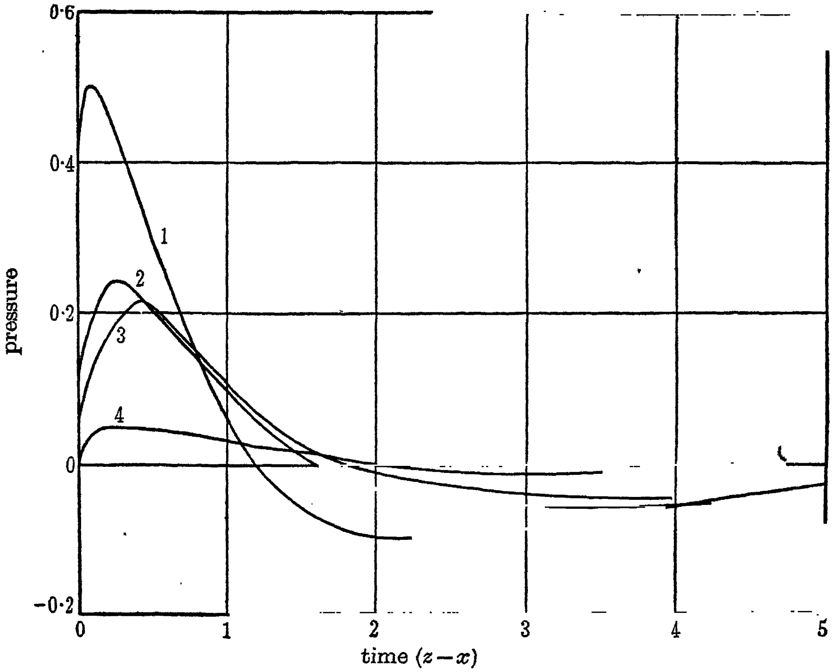


FIGURE 7. Pressure-time curves at points on the back of the wall. Curve 1, $x = 0{\cdot}05$, $y = 0$; curve 2, $x = 0{\cdot}75$, $y = 0$; curve 3, $x = 1{\cdot}5$, $y = 0$; curve 4, $x = 20$, $y = 0$.

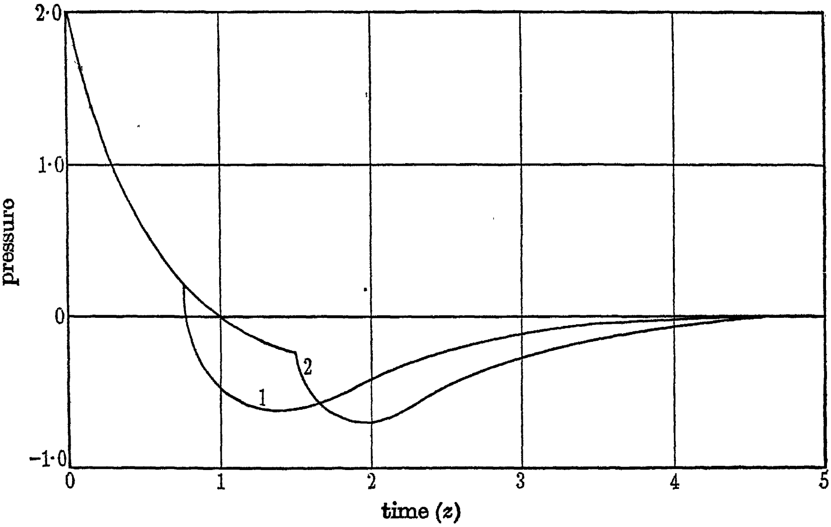


FIGURE 8. Difference of the pressures on the two sides of the wall at two points, (1) $x = 0{\cdot}75$, (2) $x = 1{\cdot}5$.

of the incident pulse until the reflected wave has arrived (simultaneously at the back and at the front of the wall); its effect appears as a considerable negative pressure difference, so that a comparatively large force acts on the wall in the direction opposite to that of the propagation of the pulse.

CALCULATIONS: VARIOUS FORMS OF INCIDENT PULSE

Several other cases in addition to that given in the last section have been discussed, though less thoroughly.

For a pulse where the pressure is given by

$$\begin{aligned} p_0(z) &= 2z, && (0 \leqslant z \leqslant \tfrac{1}{2}) \\ &= 2(1-z), && (\tfrac{1}{2} \leqslant z \leqslant 1) \\ &= 0, && (z \geqslant 1) \end{aligned}$$

the pressure distribution near the edge of the shadow has been calculated in a way analogous to the discussion in the preceding section. The function $P(X, T)$ becomes in this case

$$P = X\gamma\left(\frac{T}{X}\right) \quad (0 < T \leqslant \tfrac{1}{2}), \qquad P = X\left\{\gamma\left(\frac{T}{X}\right) - 2\gamma\left(\frac{T-\frac{1}{2}}{X}\right)\right\} \quad (\tfrac{1}{2} \leqslant T \leqslant 1),$$

$$P = X\left\{\gamma\left(\frac{T}{X}\right) + \gamma\left(\frac{T-1}{X}\right) - 2\gamma\left(\frac{T-\frac{1}{2}}{X}\right)\right\} \quad (T \geqslant 1),$$

where $$\gamma(Z) = \frac{4}{\pi}\{(1+Z)\tan^{-1}\sqrt{Z} - \sqrt{Z}\}.$$

No approximation is necessary. The calculations were carried out for $y = -10$, x ranging from -1 to $+1$. A surface whose height represents the pressure can be constructed in the same way as before, and it is shown in oblique projection in figure 9. As this pulse does not start with a discontinuous rise in pressure, the discontinuity occurring in figure 4 is not present here; the only indications of the fact that the incident pulse is being propagated outside the shadow are here the discontinuities in slope at the highest points of the pressure-time curves, and at time $\tau = 1$, outside the shadow. There is thus no marked boundary of the shadow here at all; this may be expected to apply to any pulse where the pressure rises initially at a finite rate.

When the incident pulse is 'rectangular', i.e. the pressure is given by

$$p_0(z) = 1 \quad (0 \leqslant z \leqslant 1), \qquad p_0(z) = 0 \quad (z > 1),$$

$P(X, T)$ is easily evaluated. An application of (20a) yields

$$P(X,T) = \frac{2}{\pi}\tan^{-1}\sqrt{\frac{T}{X}} \quad (T \leqslant 1)$$

$$= \frac{2}{\pi}\left(\tan^{-1}\sqrt{\frac{T}{X}} - \tan^{-1}\sqrt{\frac{T-1}{X}}\right). \quad (T \geqslant 1)$$

Some of the diffracted pressure-time curves in this case will be found in Part II of this paper.

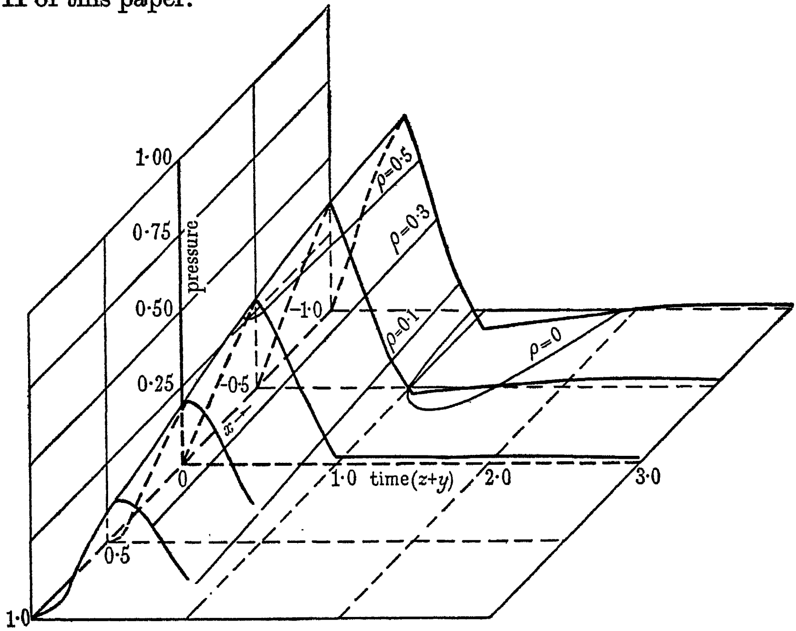


FIGURE 9. Surface $p = p(x, z)$ at $y = -10$ with level contours.

With the aid of the formulae (21), any type of incident pulse can be investigated; there is moreover no difficulty in dealing with the diffraction of a pulse whose form is only known from experiment, as it is possible to proceed by numerical integration from the start. It will be found that the case of the sharp-fronted pulse treated in the preceding section is fairly representative of a large class of pulses, viz., those which are produced suddenly and only for a short time interval. The 'rectangular pulse' is less likely to occur in practice.

In conclusion, I want to thank Professor G. I. Taylor who suggested these calculations and to whom I am indebted for much valuable advice during their progress.

REFERENCES

Bocher 1909 *Integral equations.* Camb. Math. Tracts, no. 10, p. 10.
Jahnke-Emde 1933 *Tables of Functions* (*Funktionentafeln*), 2nd ed. Leipzig, Berlin, 1933.
Lamb 1906 On Sommerfeld's diffraction problem, etc. *Proc. Lond. Math. Soc.* (2) **4**, 190.
Lamb 1910 Diffraction of a solitary wave. *Proc. Lond. Math. Soc.* (2) **8**, 422.
Sommerfeld 1895 Mathematische Theorie der Diffraktion. *Math. Ann.* **47**, 317.
Sommerfeld 1901 Theoretisches zur Beugung der Röntgenstrahlen. *Z. Math. phys.* **46**, 11.

The diffraction of sound pulses
II. Diffraction by an infinite wedge

BY F. G. FRIEDLANDER

(*Communicated by G. I. Taylor, F.R.S.—Received* 29 *February* 1940)

This part is a report on some calculations of the diffraction effects of infinite wedges which were carried out by means of a solution obtained by Sommerfeld. In all cases but one the incident pressure pulse was taken to be 'rectangular', i.e. to consist of unit pressure rise persisting for unit time; in one case a certain pulse discussed in detail in Part I of this paper was taken. It is found that the pressure at the summit of a wedge is actually greater than that of the incident pulse, but dies down in the shadow. The general character of the diffracted pressure-time curves is very similar to that of the curves obtained in the case of the half-plane.

In Part I of this paper the diffraction of plane sound pulses by a semi-infinite plane has been discussed. For purposes of comparison, some calculations have been made of the diffraction effect of an infinite wedge. The solution of this problem has been given by Sommerfeld (1901), but without any numerical application. The pressure-time curves were calculated for points on the back face of certain wedges, mainly in the case where the incident pulse is of the 'rectangular' type. The pressure at the summit of a wedge is actually bigger than the incident pressure, and in consequence it is found that the attenuation of the pressure due to a wedge is not as large as that due to a half-plane, while the shapes of the curves are very similar. In one case, the incident pulse has been taken as

$$p_0(z) = (1-z)\,e^{-z}, \quad (z > 0)$$

a case which is discussed in detail in Part I for the half-plane.

The solution given by Sommerfeld depends on the many-valued function

$$u_n(\phi') = -\frac{\sin\frac{\pi}{n}}{2\pi n} \times \int_0^\infty f(ct - r\cosh b)\left(\frac{1}{\cos\frac{\phi-\phi'+ib}{n} - \cos\frac{\pi}{n}} + \frac{1}{\cos\frac{\phi-\phi'-ib}{n} - \cos\frac{\pi}{n}}\right) db, \tag{1}$$

which has the origin ($r = 0$) as branch point, and n branches. r and ϕ being the polar co-ordinates defined by

$$x = r\cos\phi; \quad y = r\sin\phi,$$

the function (1) satisfies the sound equation

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2}. \tag{2}$$

Here Φ is the velocity potential and c the velocity of sound. As n tends to infinity, (1) tends to

$$u_\infty(\phi') = -\int_0^\infty f(ct - r\cosh b)\left(\frac{1}{\pi^2 - (\phi-\phi'+ib)^2} + \frac{1}{\pi^2 - (\phi-\phi'-ib)^2}\right) db. \tag{1a}$$

With the aid of (1) it is possible to write down the velocity potential of the motion resulting from the diffraction of a plane pulse by an infinite wedge whose angle is commensurable with π. The function f appearing in (1) is identified with the velocity potential of the incident pulse which is thus given by

$$f\{ct - r\cos(\phi - \phi')\}. \tag{3}$$

ϕ' has been taken as the angle which the direction of this incident pulse makes with that face of the wedge which is struck by it. This face is taken to be in the direction $\phi = 0$. The origin is the summit of the wedge, so that r is the distance of any point from the summit. Let $\frac{2m-n}{m}\pi$ be the angle of the wedge, so that the external angle is given by $\frac{n}{m}\pi$ where m and n are integers prime to each other. Then the velocity potential of the resulting motion is, in the shadow,

$$\left.\begin{aligned}\Phi = u_n(\phi') + u_n\left(\frac{2\pi n}{m} + \phi'\right) + u_n\left(\frac{4\pi n}{m} + \phi'\right) + \ldots + u_n\left(\frac{2(m-1)n\pi}{m} + \phi'\right) \\ + u_n(-\phi') + u_n\left(\frac{2\pi n}{m} - \phi'\right) + u_n\left(\frac{4\pi n}{m} - \phi'\right) + \ldots + u_n\left(\frac{2(m-1)n\pi}{m} - \phi'\right).\end{aligned}\right\} \tag{4}$$

Outside the shadow the transmitted and reflected waves respectively have to be added to this expression. If the angle of the wedge is not commensurable with π, the function (1a) may be employed, (4) being replaced by the infinite series

$$\left.\begin{aligned}\Phi = u_\infty(\phi')+u_\infty(2\alpha+\phi')+u_\infty(4\alpha+\phi')+\dots\\ +u_\infty(-\phi')+u_\infty(2\alpha-\phi')+u_\infty(4\alpha-\phi')+\dots,\end{aligned}\right\} \tag{4a}$$

where α is the external angle of the wedge.

To obtain the pressure, (4) must be differentiated with respect to the time, since

$$p = \rho\frac{\partial\Phi}{\partial t}.$$

It is easily seen that the result is of the same form as (4), but that the function f in (1) must now be replaced by p_0, the incident pressure pulse, so that

$$u_n(\phi') = -\frac{\sin\dfrac{\pi}{n}}{2\pi n} \times\int_0^\infty p_0(ct-r\cosh b)\left(\frac{1}{\cos\dfrac{\phi-\phi'+ib}{n}-\cos\dfrac{\pi}{n}}+\frac{1}{\cos\dfrac{\phi-\phi'-ib}{n}-\cos\dfrac{\pi}{n}}\right)db. \tag{5}$$

Similar expressions hold for u_∞. If the incident pressure pulse is 'rectangular', i.e. given by

$$p_0(z) = 0 \quad (z<0); \qquad p_0(z) = 1 \quad (0\leqslant z\leqslant 1); \qquad p_0(z) = 0 \quad (z>1),$$

then it is found that

$$u_n(\phi') = -\frac{1}{\pi}\left[\tan^{-1}\left(\frac{\sin\dfrac{\pi}{n}\sinh\dfrac{b}{n}}{\cos\dfrac{\delta}{n}-\cos\dfrac{\pi}{n}\cosh\dfrac{b}{n}}\right)\right]_{b_1}^{b_2}, \tag{6}$$

where $$\delta = \phi-\phi'$$

and the limits are

$$b_1 = 0; \quad b_2 = \cosh^{-1}\frac{ct}{r}; \quad 0\leqslant ct-r\leqslant 1,$$

$$b_1 = \cosh^{-1}\frac{ct-1}{r}; \quad b_2 = \cosh^{-1}\frac{ct}{r}; \quad ct-r\geqslant 1.$$

Similarly, $$u_\infty(\phi') = -\frac{1}{\pi}\left[\tan^{-1}\left(\frac{2\pi b}{\pi^2-b^2-\delta^2}\right)\right]_{b_1}^{b_2} \tag{6a}$$

taken between the same limits. At the summit of the wedge,

$$u_n(\phi') = \frac{1}{n} \quad (0 < ct \leqslant 1); \quad = 0 \quad (ct > 1)$$

hence it is found from (4) that the pressure there is

$$\left.\begin{aligned} p(r=0) &= \frac{2m}{n} = \frac{2}{k}, \quad (0 \leqslant ct \leqslant 1) \\ &= 0 \quad \ldots, \quad (ct > 1) \end{aligned}\right\} \tag{7}$$

where $k\pi$ is the external angle of the wedge. Thus the pressure there is equal to the incident pressure only in the case of the half-plane; while for the limiting case $k = 1$ it is doubled, which is of course a well-known property of reflexion. It is clear by continuity that this result must also hold for a wedge whose angle is not commensurable with π.

With the aid of formula (6) some calculations of pressure-time curves have been made. The point at which the pressure has been calculated has in all cases been taken on the back of the wedge (i.e. on the wedge, but in the 'shadow') at unit distance from the summit of the wedge. The wedges, and 'angles of incidence', for which the calculations have been made, are as follows:

(1) 120°-wedge; $\phi' = 30°$
(2) 60°-wedge; position 1, $\phi' = 60°$
(3) 60°-wedge; position 2, $\phi' = 90°$
(4) 60°-wedge; position 3, $\phi' = 30°$
(5) 90°-wedge; $\phi' = 0°$

(cf. figure 1). The pressure-time curve at a point at unit distance below the top of a half-plane, at the back, has also been calculated (case (6)).

Figure 2 shows the curves corresponding to the cases (1), (2) and (6)—that is to say, all the cases in which the direction of the incident pulse is perpendicular to the plane of symmetry of the wedge. The maximum pressure in case (1) is still well above 1. Owing to the simple form of (6) it is easy to calculate the maximum pressure at any point on the back face of the wedge, and figure 3 shows these maximum pressures for the same wedges and incidences as appear in figure 2. It is seen that the falling off of pressure with increasing distance from the summit at the back of the 120°-wedge is fairly rapid.

Figure 4 shows the three curves obtained for the 60°-wedge (cases (2), (3) and (4)). It is seen that the greatest attenuation of pressure occurs when the back of the wedge is perpendicular to the direction of the incident pulse. In figure 5 cases (5), (4) and (6) are compared; in each case the direction of the incident pulse is perpendicular to the back face of the wedge.

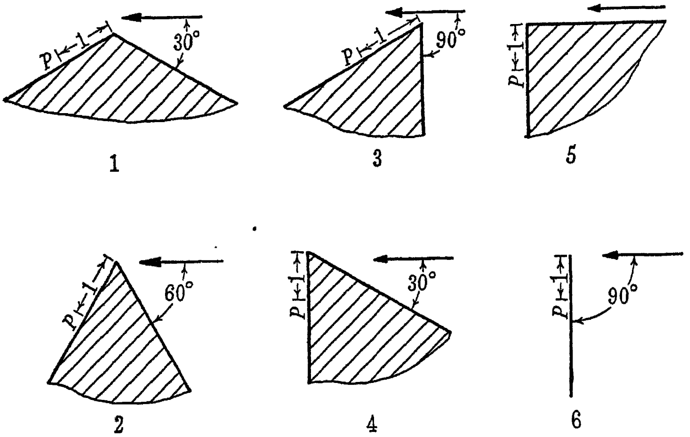


Figure 1. Pressure-time curves have been calculated at the points marked P in the figure. The arrow in each case denotes the direction of the incident pulse.

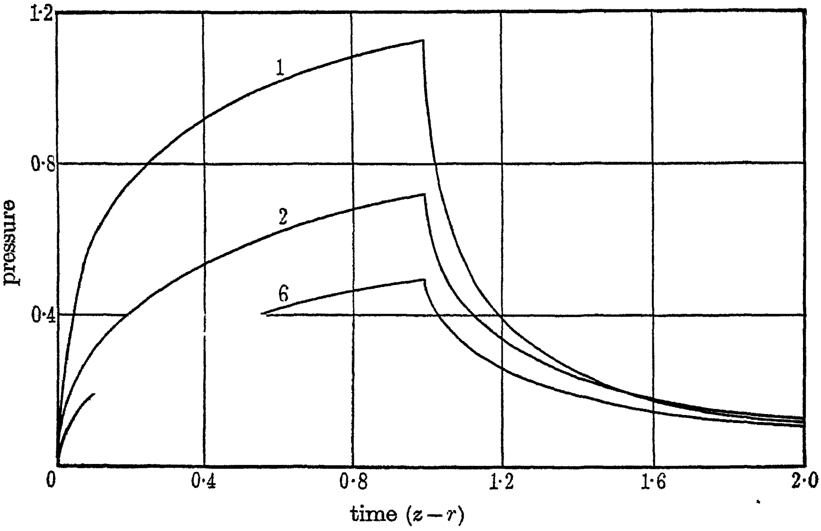


Figure 2. Direction of incident pulse perpendicular to the plane of symmetry of the wedge: (1) 120°-wedge, (2) 60°-wedge, (3) half-plane.

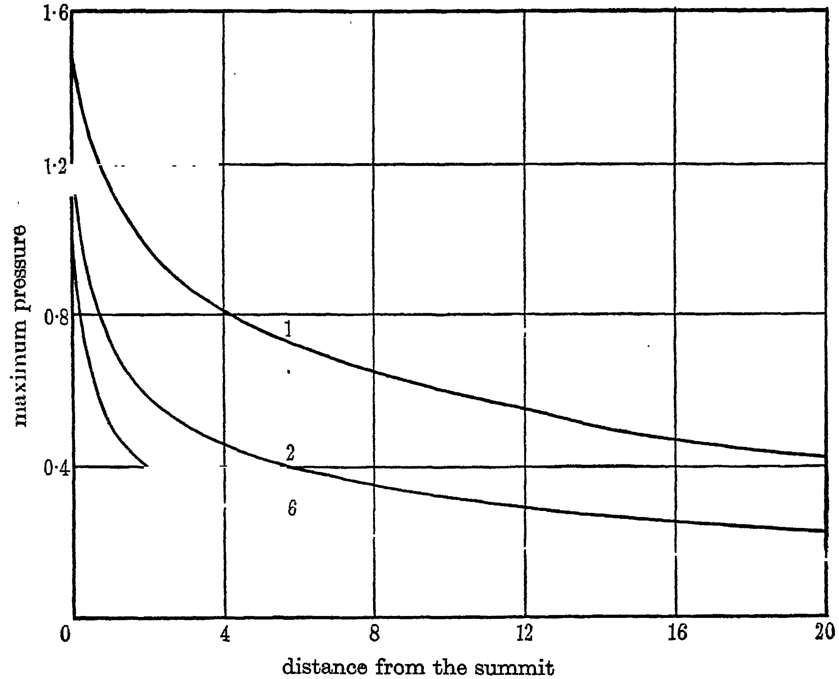


FIGURE 3. Maximum pressure as a function of distance from the summit on the back face of wedges: (1) 120°-wedge, (2) 60°-wedge, position 1, (3) half-plane.

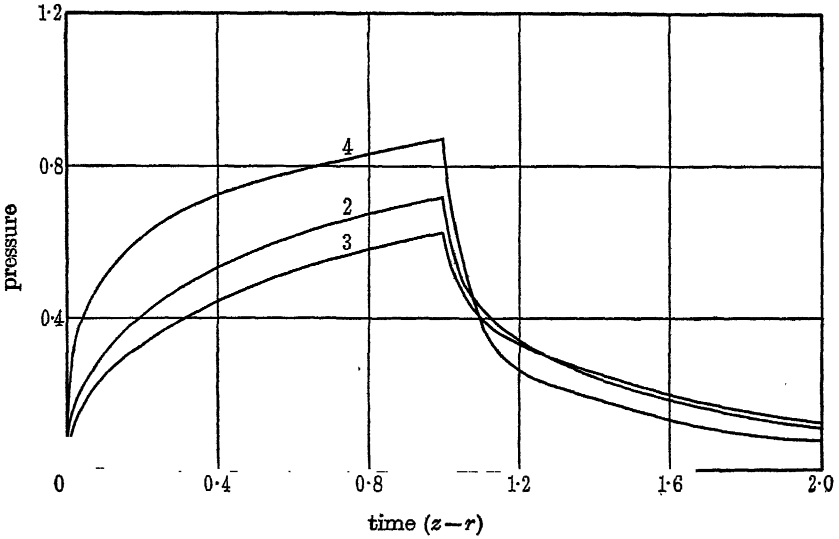


FIGURE 4. Three different positions of a 60°-wedge, (2) position 1, (3) position 2, (4) position 3.

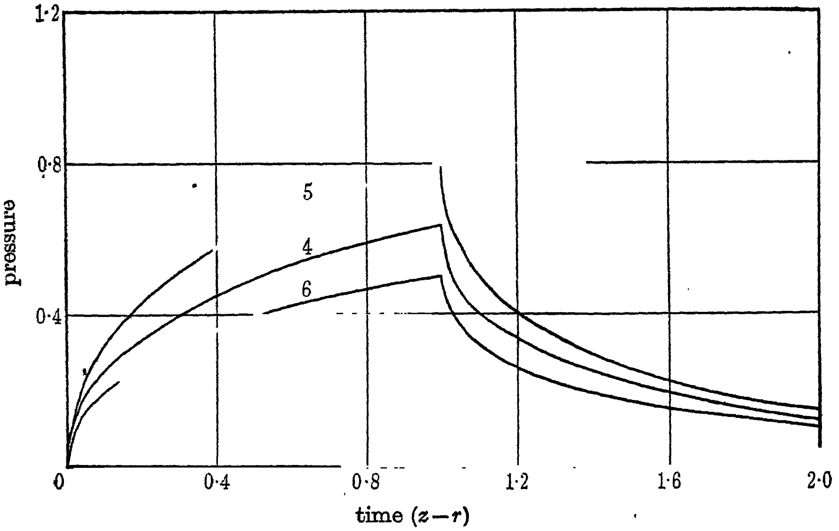


Figure 5. Direction of incident pulse perpendicular to the back face of the wedge: (4) 60°-wedge, position 3, (5) 30°-wedge, (6) half-plane.

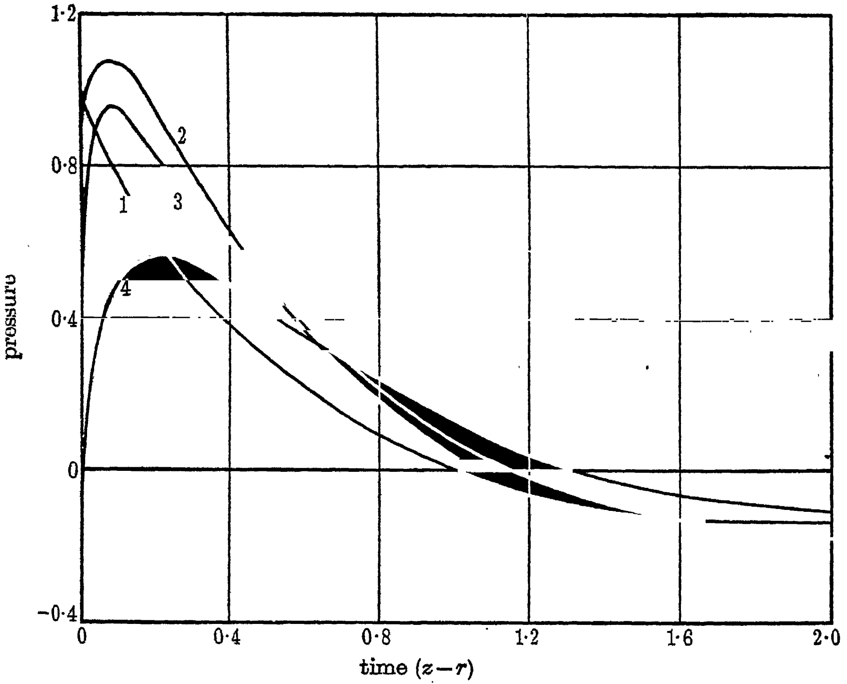


Figure 6. Diffraction of $P_0 = (1-z)\,e^{-z}$ by a 120° wedge: pressure-time curves for points on the back face: (1) incident pulse, (2) $r = 0{\cdot}05$, $\phi = 240°$, (3) $r = 0{\cdot}10$, $\phi = 240°$, (4) $r = 1{\cdot}0$, $\phi = 240°$.

In the case of a 120°-wedge whose plane of symmetry is at right angles to the direction of the incident pulse (case (1)), some diffracted pressure-time curves at points on the back of the wedge have been found for the incident pressure pulse

$$p_0(z) = 0, \quad (z < 0)$$
$$p_0(z) = (1-z)\,e^{-z}, \quad (z \geqslant 0)$$

by numerical integration. They are shown in figure 6. The pressure at the summit is still 1·5 times as great as the incident pressure, but the curve for the point $r = 0{\cdot}05$ shows that the maximum of the pressure reached at any point decreases quickly with increasing distance from the summit. It will be noticed that the shapes of the curves obtained are very much like those obtained in the case of a half-plane.

TABLE 1. RECTANGULAR PRESSURE PULSE: PRESSURES AT POINTS ON THE BACK FACE OF WEDGES, AT UNIT DISTANCE BELOW THE SUMMIT

$ct-r$	1	2	3	4	5	6
0·0	0·000	0·000	0·000	0·000	0·000	0·000
0·2	0·758	0·407	0·628	0·337	0·434	0·268
0·4	0·924	0·538	0·726	0·452	0·576	0·359
0·6	1·016	0·621	0·792	0·530	0·670	0·420
0·8	1·089	0·676	0·836	0·589	0·740	0·465
1·0	1·124	0·719	0·868	0·633	0·790	0·500
1·2	0·394	0·343	0·264	0·334	0·398	0·261
1·4	0·260	0·241	0·186	0·248	0·291	0·194
1·6	0·193	0·182	0·134	0·197	0·227	0·155
1·8	0·139	0·141	0·097	0·150	0·181	0·127
2·0	0·120	0·128	0·080	0·127	0·152	0·108

REFERENCE

Sommerfeld 1901 *Z. Math. Phys.* **46**, 11. See also a paper by Carslaw, *Proc. Lond. Math. Soc.* **30** (1899) 121.

The diffraction of sound pulses

III. Note on an integral occurring in the theory of diffraction by a semi-infinite screen

By F. G. Friedlander

(*Communicated by G. I. Taylor, F.R.S.—Received* 26 *November* 1941)

This note completes the development of certain approximations of the integral $P(X, T)$ which was defined in the first paper of this series. The two expansions for small and large X respectively whose first terms were then given are stated in full and the magnitude of the remainders estimated. Another expansion of $P(X, T)$ is proposed which is valid when T, as well as X, is small.

In the first of these papers on the diffractions of sound pulses* (which will in the following be referred to as 'Diffraction I') Sommerfeld's theory of the diffraction of plane pulses by a semi-infinite screen was discussed and the solution was expressed in terms of the integral

$$P(X,T)=\frac{\sqrt{X}}{\pi}\int_0^T\frac{p_0(u)\,du}{(T+X-u)\sqrt{(T-u)}} \tag{1}$$

('Diffraction I', equation (20a)). Making the substitution

$$\zeta^2=T-u, \tag{2}$$

this can be written

$$P(X,T)=\frac{2\sqrt{X}}{\pi}\int_0^{\sqrt{T}}\frac{p_0(T-\zeta^2)\,d\zeta}{X+\zeta^2}. \tag{3}$$

For small and large values of X respectively, certain approximate expressions of $P(X, T)$ were obtained, but only incomplete estimates of the accuracy of these approximations were given. It is proposed to complete the development of these approximations in the present note.

The approximation for small values of X was obtained as follows: writing

$$\begin{aligned}P(X,T)&=\int_0^{T+X}\frac{\phi(u)\,du}{\sqrt{(T+X-u)}}-\int_T^{T+X}\frac{\phi(u)\,du}{\sqrt{(T+X-u)}}\\&=p_0(T+X)-\int_T^{T+X}\frac{\phi(u)\,du}{\sqrt{(T+X-u)}},\end{aligned}$$

---

* See p. 322.

the second term on the right-hand side was transformed by partial integration,

$$\int_T^{T+X} \frac{\phi(u)\,du}{\sqrt{(T+X-u)}} = 2\phi(T)\sqrt{X} + 2\int_T^{T+X} \phi'(u)\sqrt{(T+X-u)}\,du.$$

If the integral on the right-hand side of this equation is again subjected to partial integration, and the process continued, the result after n partial integrations will be

$$\int_T^{T+X} \frac{\phi(u)\,du}{\sqrt{(T+X-u)}} = \frac{2}{1}\phi(T)\,X^{\frac{1}{2}} + \frac{2}{1}\cdot\frac{2}{3}\phi'(T)\,X^{\frac{3}{2}} + \dots$$
$$+ \frac{2^n X^{n-\frac{1}{2}}}{1.3\dots(2n-1)}\phi^{(n-1)}(T) + \frac{2^n}{1.3\dots(2n-1)}\int_T^{T+X} \phi^{(n)}(u)\,(T+X-u)^{n-\frac{1}{2}}\,du. \tag{4}$$

Making use of the mean value theorem, we write the remainder, i.e. the integral on the right-hand side of (4), in the form

$$\frac{2^{n+1}X^{n+\frac{1}{2}}}{1.3\dots(2n+1)}\phi^{(n)}(T+\theta X) \quad (0 \leqslant \theta \leqslant 1). \tag{5}$$

Now (by 'Diffraction I', equation (16*a*))

$$\phi^{(n)}(u) = \left(\frac{d}{du}\right)^{n+1} \frac{1}{\pi}\int_0^u \frac{p_0(z)\,dz}{\sqrt{(u-z)}} = \frac{1}{\pi}\left\{\frac{(-1)^n\,1.3\dots(2n-1)}{2^n u^{n+\frac{1}{2}}}p_0(0)\right.$$
$$\left. + \frac{(-1)^{n-1}\,1.3\dots(2n-1)}{2^{n-1}u^{n-\frac{1}{2}}}p_0'(0) + \dots + \frac{p_0^{(n)}(0)}{u^{\frac{1}{2}}} + 2\int_0^{\sqrt{u}} p_0^{(n+1)}(u-\zeta^2)\,d\zeta\right\}.$$

Hence it is found that

$$|\phi^{(n)}(T+\theta X)| \leqq \frac{1}{\pi}\left\{\frac{1.3\dots(2n-1)}{2^n T^{n+\frac{1}{2}}}\,|p_0(0)| + \dots\right.$$
$$\left. + \frac{|p_0^{(n)}(0)|}{T^{\frac{1}{2}}} + 2(T+X)^{\frac{1}{2}}A_{n+1}\right\}; \tag{5a}$$

where A_{n+1} is the upper bound of $|p_0^{(n+1)}(u)|$ in the interval $(0, T+X)$.

This expansion fails when T, as well as X, is small. To overcome this difficulty, another series expansion of $P(X, T)$ can be proposed.

Consider the function

$$Q(X,T) = \int_0^{\sqrt{T}} \frac{p_0(T+X) - p_0(T-\zeta^2)}{X+\zeta^2}\,d\zeta \tag{6}$$

or

$$Q(X,T) = \frac{\pi}{2\sqrt{X}}\left\{\frac{2}{\pi}p_0(T+X)\tan^{-1}\sqrt{\frac{T}{X}} - P(X,T)\right\}. \tag{6a}$$

This function and its derivatives with respect to X are regular when $X = 0$. It can be expanded as a Taylor series in powers of X.

Differentiating under the integral sign in (6) it is found that

$$\frac{\partial^m Q(X,T)}{\partial X^m} = \int_0^{\sqrt{T}} d\zeta \left\{ \frac{p_0^{(m)}(T+X)}{X+\zeta^2} - 1!\binom{m}{1} \frac{p_0^{(m-1)}(T+X)}{(X+\zeta^2)^2} \right.$$
$$\left. + 2!\binom{m}{2} \frac{p_0^{(m-2)}(T+X)}{(X+\zeta^2)^3} - \ldots + (-1)^m \frac{m!}{(X+\zeta^2)^{m+1}} [p_0(T+X) - p_0(T-\zeta^2)] \right\},$$

which can be written

$$\frac{\partial^m Q(X,T)}{\partial X^m} = \int_0^{\sqrt{T}} \frac{(-1)^{m+1} m!\, d\zeta}{(X+\zeta^2)^{m+1}} \left\{ p_0(T-\zeta^2) - p_0(T+X) \right.$$
$$\left. + \frac{(X+\zeta^2)}{1!} p_0'(T+X) - \ldots + (-1)^{m+1} \frac{(X+\zeta^2)^m}{m!} p_0^{(m)}(T+X) \right\}.$$

Now the expression inside the curved brackets in the integrand on the right-hand side of the last equation is the remainder of a Taylor series which can be expressed as a definite integral; this integral can be obtained as follows: put

$$R_{m+1} = p_0(T-\zeta^2) - p_0(T+X) + (X+\zeta^2)\, p_0'(T+X) - \ldots$$
$$+ (-1)^{m+1} \frac{(X+\zeta^2)^m}{m!} p_0^{(m)}(T+X); \qquad (7)$$

differentiating, $\dfrac{\partial R_{m+1}}{\partial X} = (-1)^{m+1} \dfrac{(X+\zeta^2)^m}{m!} p_0^{(m+1)}(T+X)$,

as all terms except the last one cancel. Hence

$$R_{m+1} = \frac{(-1)^{m+1}}{m!} \int_{-\zeta^2}^{X} (\zeta^2+z)^m p_0^{(m+1)}(T+z)\, dz.$$

If a new variable of integration is introduced by the substitution

$$z = X - (X+\zeta^2)\, t,$$

this becomes

$$R_{m+1} = \frac{(-1)^{m+1} (X+\zeta^2)^{m+1}}{m!} \int_0^1 (1-t)^m p_0^{(m+1)}\{T + X - (X+\zeta^2)\, t\}\, dt.$$

Substituting, the mth derivative of $Q(X,T)$ with respect to X is obtained in the form

$$\frac{\partial^m Q(X,T)}{\partial X^m} = \int_0^{\sqrt{T}} d\zeta \int_0^1 (1-t)^m p_0^{(m+1)}\{T + X - (X+\zeta^2)\, t\}\, dt. \qquad (8)$$

The Taylor series expansion of $Q(X, T)$ can now be written down. Using Lagrange's form of the remainder,

$$Q(X,T) = Q(0,T) + X\left(\frac{\partial Q}{\partial X}\right)_{X=0} + \ldots + \frac{X^{n-1}}{(n-1)!}\left(\frac{\partial^{n-1} Q}{\partial X^{n-1}}\right)_{X=0} + \frac{X^n}{n!}\left(\frac{\partial^n Q}{\partial X^n}\right)_{X=\theta X}, \tag{9}$$

and substituting this expansion in (6*a*),

$$\begin{aligned} P(X,T) &= \frac{2}{\pi} p_0(T+X)\tan^{-1}\sqrt{\frac{T}{X}} \\ &\quad - \frac{2\sqrt{X}}{\pi}\sum_{m=0}^{n-1}\frac{X^m}{m!}\int_0^{\sqrt{T}} d\zeta \int_0^1 (1-t)^m p_0^{(m+1)}(T - t\zeta^2)\,dt \\ &\quad - \frac{2X^{n+\frac{1}{2}}}{n!}\int_0^{\sqrt{T}} d\zeta \int_0^1 (1-t)^n p_0^{(n+1)}\{T + \theta X - (\theta X + \zeta^2)\,t\}\,dt. \end{aligned} \tag{10}$$

An upper limit for the absolute magnitude of the remainder can then be easily given; it is

$$2A_{n+1}\sqrt{(XT)}\frac{X^n}{n!}, \tag{11}$$

where A_{n+1} has the same meaning as in equation (5*a*).

Finally, consider the expansion of $P(X, T)$ for large values of X, the first term of which was given in 'Diffraction I', equation (22).

Writing
$$P(X,T) = \frac{\sqrt{X}}{\pi}\int_0^T \frac{p_0(T-u)\,du}{(X+u)\sqrt{u}}, \tag{12}$$

and assuming that $T < X$, substitute the expansion

$$\frac{1}{X+u} = \frac{1}{X}\left\{1 - \frac{u}{X} + \frac{u^2}{X^2} - \ldots + (-1)^{n-1}\frac{u^{n-1}}{X^{n-1}} + (-1)^n \frac{(u/X)^n}{1+u/X}\right\}.$$

Then

$$\begin{aligned} P(X,T) &= \frac{1}{\pi\sqrt{X}}\sum_{m=0}^{n-1}\frac{(-1)^m}{X^m}\int_0^T u^{m-\frac{1}{2}} p_0(T-u)\,du \\ &\qquad\qquad + \frac{(-1)^n}{\pi X^{n-\frac{1}{2}}}\int_0^T \frac{u^{n-\frac{1}{2}} p_0(T-u)\,du}{X+u}. \end{aligned} \tag{13}$$

An obvious upper bound of the absolute value of the remainder is

$$\left|\frac{(-1)^n}{\pi X^{n-\frac{1}{2}}}\int_0^T \frac{u^{n-\frac{1}{2}} p_0(T-u)}{X+u}\,du\right| \leqq \frac{K}{\pi}\left(\frac{T}{X}\right)^{n+\frac{1}{2}}, \tag{14}$$

where
$$|\,p_0(u)\,| \leqq K \quad (0 \leqq u \leqq T).$$

The diffraction of sound pulses
IV. On a paradox in the theory of reflexion

By F. G. Friedlander

(*Communicated by G. I. Taylor, F.R.S.—Received* 26 *November* 1941)

Sommerfeld's theory of diffraction, which has been discussed in detail in the preceding papers, is here applied to the elucidation of a certain paradox—the doubling of pressure at an infinite reflecting plane parallel to which a pulse is travelling, which is obtained if this is considered as a limiting case of reflexion. It is suggested that this paradox is an indication that the simple theory of the reflexion of plane pulses by an infinite plane, which is in any case only an approximation applying to parts of a reflector far removed from the edges, breaks down when the angle between the direction of propagation of the incident pulse and the reflector becomes small. A detailed examination of the reflexion of plane pulses by a semi-infinite screen shows that when this angle becomes small, the diffracted pulse which emanates from the edge of the screen must be taken into account, so that the region in which the simple theory holds as an approximation moves away from the edge. In the limit, when the incident pulse travels parallel to the screen, the simple theory must be rejected at all points at a finite distance from the edge. Further examples of this mechanism are given by the reflexion of pulses by infinite wedges, some cases of which are also considered.

The simple theory of the reflexion of plane sound waves by an infinite plane reflector leads to a paradoxical conclusion when the angle of incidence, i.e. the angle between the wave front of the incident wave and the reflecting plane, tends to 90°.

The reflected wave is another plane wave of the same form as the incident one, the two wave fronts being equally inclined to the plane of the reflector; the pressure at any point is the sum of the pressures in the two waves, so that at the reflector it is double the pressure in the incident wave, irrespective of the angle of incidence. As this tends to 90°, the direction of propagation of the incident wave tends to become parallel to the reflector, but reflexion and doubling persist, and in the limit it follows from this analysis that the pressure at the reflector is doubled even when the incident wave travels along it—but then there should be no reflexion, and hence no doubling of the pressure.

As an infinite plane mirror is a mathematical conception which cannot be realized experimentally, the simple theory of reflexion, from the point of view of physics, should merely be an approximation to the rigorous theory of a plane reflector which holds at points far from its edges. The paradox which has been set out above makes it seem likely that even this interpretation is incorrect when the angle of incidence is nearly a right angle; that then the fact that a real reflector is of finite extent cannot be ignored any longer and that the diffraction of the incident wave by the edges of the reflector must be taken into account.

This will in general be a difficult mathematical problem and not even approximate methods will be available, as I propose to deal with pulses rather than with infinite trains of harmonic waves. But a certain amount of information can be extracted from the rigorous solutions of the problems of the diffraction of plane waves by a semi-infinite screen, and by wedges of any angle, which have been obtained by Sommerfeld. Only two-dimensional wave motions are considered, but the presence of one edge introduces diffracted waves which can be calculated, and the results will serve to give a more detailed picture of the type of mechanism which ensures that in the case of reflectors of finite extent the paradoxical conclusion drawn from the 'simple theory' is invalidated.

REFLEXION BY A SEMI-INFINITE SCREEN

1. I begin by considering the reflexion and diffraction of a plane wave or pulse whose wave front is parallel to the straight edge of a reflecting semi-infinite plane screen.

Let r, θ be polar co-ordinates in one of the plane sections at right angles to the plane containing the screen, the origin being taken at the point where the edge of the screen meets this section, and the angle θ being counted from that side of the screen which is struck by the incident pulse. The screen is then given by $\theta = 0$ (and, in the 'shadow', $\theta = 2\pi$); the angle of incidence will be denoted by θ'', and the angle between the direction of propagation of the incident wave and the screen by θ', so that $\theta' = \theta'' + \frac{1}{2}\pi$, and we assume that $\frac{1}{2}\pi \leqq \theta' \leqq \pi$.

If the incident wave is a pulse with a sharp wave front which reaches the edge of the screen at time $t = 0$, it follows from Sommerfeld's solution that the resulting state at time t can be considered as made up of three parts:

(i) the incident or 'transmitted' wave, whose wave front is given, in figure 1, by the two straight-line segments PP' and SS', each extending to infinity;

(ii) the reflected wave, determined in accordance with the 'simple theory', whose wave front is the (finite) straight-line segment PR, and which must be added to (i) in the triangle PRO;

(iii) the diffracted wave, with a cylindrical wave front appearing in figure 1 as the circle $RQSR$ with centre 0, and of radius $r = ct$ (where c is the velocity of sound), touching PR and SS'.

The boundary of the region in which there is disturbance is thus the broken line $P'PQSS'$; the 'simple theory' holds in the region bounded by PQ, PR and the circular arc QR.

Hence at the point P situated on that side of the screen which is struck by the incident wave, pressures in accordance with the 'simple theory' are experienced during a certain time interval after the arrival of the incident wave. Then the diffracted wave interferes with the 'simple theory' wave, and it will appear that, if the incident wave or pulse consists of excess pressure, the diffracted one will introduce suction, or rather tend to reduce the doubled pressure predicted by the simple theory.

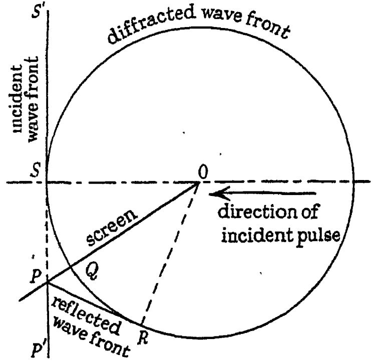


FIGURE 1. The wave fronts in the case of the diffraction of a plane pulse by a semi-infinite screen.

2. To make this point clearer, an incident wave consisting of a 'square-topped' or 'rectangular' pressure pulse will be considered; i.e. the pressure in excess of atmospheric pressure in the incident wave rises discontinuously, remains constant for a certain time interval and then falls, again discontinuously, to zero.

In order to simplify the notation, non-dimensional units will be employed from now on. The length of the rectangular pulse will be taken as the unit of length; the time will be replaced by the variable

$$\tau = ct, \tag{1}$$

which is of the dimension of a length; the excess pressure in the rectangular pulse will be adopted as the unit of pressure.

At the point P, which is at distance r from the edge of the screen, the incident (and reflected) wave arrives when

$$\tau = -r\cos\theta' \quad (\cos\theta' \leqq 0),$$

and the diffracted wave arrives when

$$\tau = r;$$

thus the 'time lag' between these waves, expressed as a corresponding distance, is

$$L = 2r\cos^2\tfrac{1}{2}\theta. \tag{2}$$

The pressure at P is found with the aid of Sommerfeld's many-valued solution of the equation of wave motion.* It is given by

$$p = [2] + u_2(\theta') + u_2(-\theta'),$$

where p is the pressure, and u_2 Sommerfeld's function for this particular case, and where the term enclosed in square brackets must be omitted unless

$$-r\cos\theta' < \tau < 1 - r\cos\theta'.$$

Further,

$$u_2(\theta') = -\frac{1}{\pi}\left[\tan^{-1}\frac{\sinh\frac{1}{2}b}{\cos\frac{1}{2}\theta'}\right]_{b_1}^{b_2},$$

where

$$b_2 = \cosh^{-1}\frac{\tau}{r}; \qquad \begin{array}{ll} b_1 = 0 & \text{if} \quad \tau - r \leqq 1, \\ b_1 = \cosh^{-1}\dfrac{\tau-1}{r} & \text{if} \quad \tau - r \geqq 1; \end{array}$$

hence it is found that the formula for p can be written

$$p = [2] - \frac{2}{\pi}\left\{\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)} - \tan^{-1}\sqrt{\left(\frac{\tau-r-1}{L}\right)}\right\}, \tag{3}$$

with the convention that if a term becomes imaginary, it is to be omitted. In detail, (3) splits up into the following set of formulae:

$L<1$		L	
$p=2$	$(-r\cos\theta'<\tau<r)$	$p=2$	$(-r\cos\theta'<\tau<1-r\cos\theta')$
$p=2-\frac{2}{\pi}\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}$	$(r\leqq\tau<1-r\cos\theta')$	$p=0$	$(1-r\cos\theta'<\tau\leqq r)$
$p=-\frac{2}{\pi}\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}$	$(1-r\cos\theta'<\tau\leqq r+1)$	$p=-\frac{2}{\pi}\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}$	$(r\leqq\tau\leqq r+1)$
$p=-\frac{2}{\pi}\left\{\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}-\tan^{-1}\sqrt{\left(\frac{\tau-r-1}{L}\right)}\right\}$	$(\tau\geqq r+1)$	$p=-\frac{2}{\pi}\left\{\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}-\tan^{-1}\sqrt{\left(\frac{\tau-r-1}{L}\right)}\right\}$	$(\tau\geqq r+1)$

$$(r = L - r\cos\theta')$$

---

* Sommerfeld, *Z. Math. Phys.* **46**, 11. See also p. 344 of this volume, where explicit formulae for the pressure are given.

These expressions show that the pressure at any point on the screen depends only on the time counted from the onset of the incident wave, and on L. I shall therefore consider in the first place only the changes in the shape of the pressure-time curve when L is varied, and afterwards discuss the bearing of these results on the changes due to a variation of θ'.

3. The doubled rectangular pulse predicted by the simple theory can always be considered to be present, but when $L < 1$ it interferes with the diffracted wave.

The effect of this interference is to eliminate 'doubling' almost completely when L is small; in fact when L (as well as r) is zero, only the pressure due to the incident wave is experienced.

As L increases, the reduction of pressure due to this interference decreases, and when L is equal to unity the diffracted wave separates as a 'tail' from the 'simple-theory' rectangular pulse. When L is increased further, the maximum suction

$$p_m = -\frac{2}{\pi}\tan^{-1}\frac{1}{\sqrt{L}} \tag{4}$$

in the 'tail' decreases; some numerical values are given in the following table:

L	1	5	20	40
$-p_m$	0·5	0·268	0·140	0·100

Hence it seems reasonable to say that when L is sufficiently large, the 'tail' may be neglected and the simple theory assumed to hold; but it is important to note that although the value reached by the (negative) excess pressure in the tail become negligible, the area between it and the axis, i.e. the 'total impulse' of the tail, remains finite and constant (and equal to the total impulse of the incident pulse).

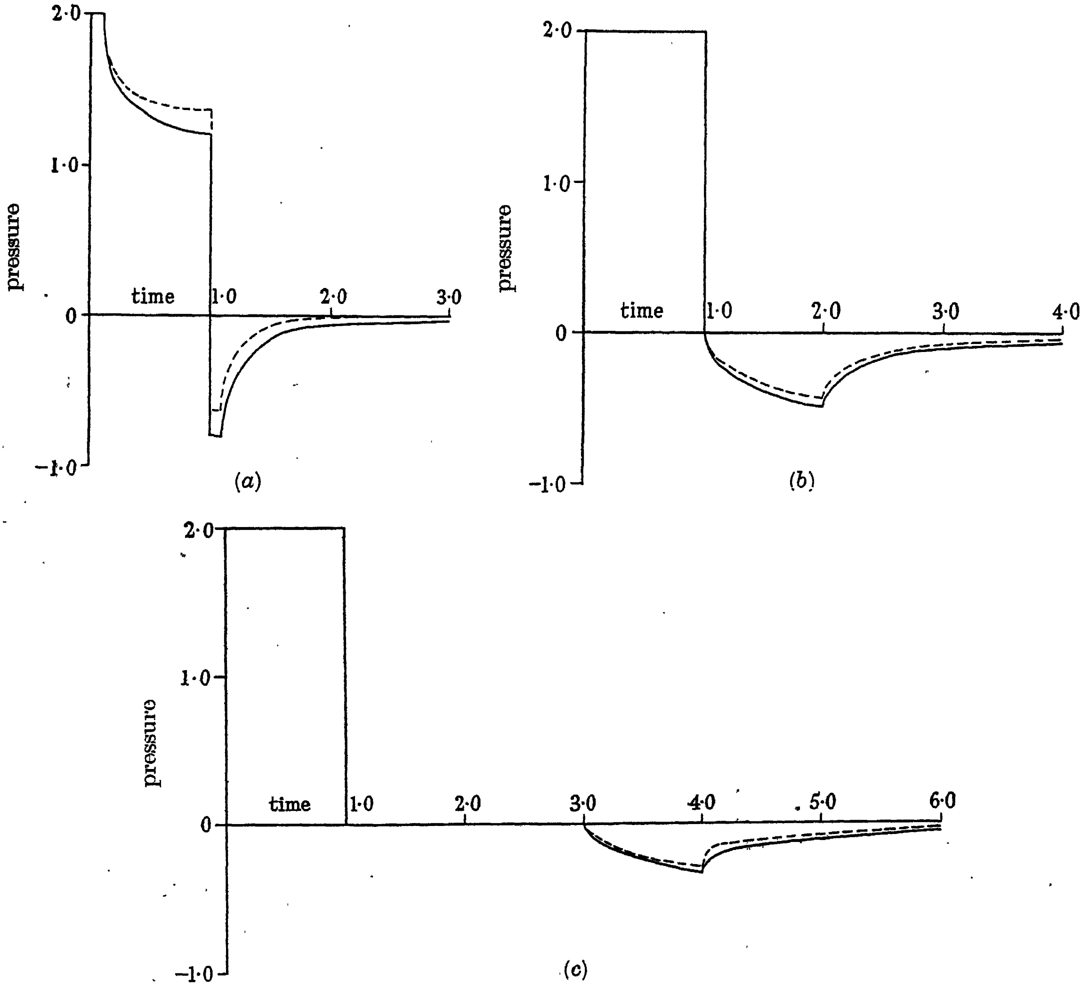
In figure 2 three pressure-time curves are shown to illustrate these considerations.

When θ' is varied, the value of r for which L assumes some specific value also varies. If L is fixed, r will become large when θ' approaches 180°; e.g. for r_0, the value of r for which $L = 1$, so that the 'tail' just separates from the rectangular pulse, some numerical values are shown in the following table:

θ'	90°	105°	120°	135°	150°	165°	180°
r_0	1·00	1·35	2·00	3·41	7·46	29·3	∞

The region in which the simple theory holds as an approximation extends from some value of r, which certainly exceeds r_0, to infinity. As θ' approaches 180°, i.e. as the direction of propagation of the incident wave tends to become parallel to the screen, this region moves away from the origin, and in the limit the 'simple theory' must be rejected at all points on the screen at a finite distance from the edge.

Thus the paradox has disappeared, or rather it has been moved to infinity; but this agrees with its original derivation for an infinite plane reflector, where it was assumed that there was an infinite length of reflector in front of the point considered.



FIGURES 2*a*–2*c*. Pressure-time curves: (*a*) $L=0{\cdot}1$; (*b*) $L=1{\cdot}0$; (*c*) $L=3$
——— half-plane - - - - - 90°-wedge

The pressure-time curves for the 30°-wedge are indistinguishable from those for the half-plane on this scale.

4. Consider now an incident pulse of arbitrary shape, whose pressure is given by

$$p_0(\tau + r\cos\theta'), \tag{5}$$

where $p_0(\tau + r\cos\theta') = 0$ if $\tau + r\cos\theta' < 0$. Using the general expression for u_2 (see p. 346), it is found, after some transformation, that

$$p = 2p_0(\tau + r\cos\theta') - \frac{\sqrt{L}}{\pi}\int_0^{\tau-r} p_0(\tau - r - u)\frac{du}{(L+u)\sqrt{u}}. \tag{6}$$

The second term on the right-hand side of this equation has an interesting physical meaning: it can be written

$$-\frac{\sqrt{L}}{\pi}\int_0^{\tau-r} p_0(\tau - r - u)\frac{du}{(L+u)\sqrt{u}} = \int_0^{\tau-r} p_0(u)\frac{d}{du}\left(\frac{2}{\pi}\tan^{-1}\sqrt{\frac{\tau-r-u}{L}}\right)du. \tag{6 a}$$

Now divide the interval $(0, \tau - r)$ into a number of small intervals, and consider the sum

$$-\sum_{k=0}^{n} p_0(u_k)\frac{2}{\pi}\left\{\tan^{-1}\sqrt{\frac{\tau-r-u_k}{L}} - \tan^{-1}\sqrt{\frac{\tau-r-u_{k+1}}{L}}\right\}$$

$$(0 = u_0 < u_1 < u_2 < \ldots < u_{n-1} < u_n = \tau - r).$$

In the limit, when $n \to \infty$ and the intervals of the subdivision become infinitely small, it tends, by the definition of an integral, to the integral (6a); but each term of this sum is the diffracted wave due to an 'elementary' rectangular pulse, of 'height' $p_0(u_k)$ and of length $u_{k+1} - u_k$. That is to say, the integral on the right-hand side of (6) can be considered as a representation of the diffracted pulse as the sum of superposed diffracted elementary rectangular pulses.

The equation (6) includes, as a special case, that of $\theta' = 90°$ which has been discussed in detail in another paper; the integral which appears in it, except for notation and a change of the variable of integration, is identical with the function $P(X, T)$ which was then introduced.* It follows from the approximate expressions derived for this function that it becomes very small when L is large. Taken together with the fact that there is here, as before, a 'time lag' L between the direct and the diffracted pulses, it can again be said that the simple theory, which treats the reflector as of infinite extent, holds as an approximation at points sufficiently far from the edge of the screen. It must be noted, however, that it can again be proved that the total impulse of the diffracted wave is constant, and numerically equal to the total impulse of the incident wave.

---

* See p. 330, equation (20a); see also the discussion of the various series developments of $P(X, T)$, p. 352 *et seq.*

When L is very small, 'doubling' disappears almost entirely. At the edge of the screen the pressure experienced is again equal to the pressure in the incident pulse; at points near to the edge the doubling persists for a short time, and the diffracted wave then quickly builds up a negative pressure which reduces the doubled pressure of the incident wave by an amount approximately equal to

$$-p_0(\tau + r\cos\theta')$$

(neglecting terms in $\sqrt{L}$ and of higher order).

The extent of the region in which the 'simple theory' holds approximately depends on the shape of the incident pulse. There is no suitable general definition of the 'length' of a pulse, and in consequence a 'simple theory' region can only be defined with respect to a given incident pulse.

5. It is convenient to have a quantity whose magnitude is a measure of the extent of the interference between the doubled incident pulse and the diffracted pulse. The integral

$$J = \int_{p \geqq 0} p\, d\tau \tag{7}$$

can be used for this purpose; it will be referred to as the 'positive impulse'. It is the contribution of the excess pressure to the total impulse, which is defined as the integral

$$I = \int_{-\infty}^{\infty} p\, d\tau.$$

An important property of the positive impulse is that when the point considered on the screen is moved to infinity, its value tends to twice the total impulse of an incident pulse involving (positive) excess pressure only.

In fact, the integral J is made up of two parts. The first,

$$J_1 = 2\int_{-r\cos\theta'}^{L - r\cos\theta'} p_0(\tau + r\cos\theta')\, d\tau = 2\int_0^L p_0(v)\, dv,$$

tends to $2I_0$, where I_0 is the total impulse of the incident pulse, as the upper limit tends to infinity (the convergence of the integral has, of course, already been tacitly assumed). Hence, if L is made large enough,

$$|\, 2I_0 - J_1 \,| \leqq \epsilon,$$

where ϵ is some arbitrary small quantity.

The second part, J_2, is the integral over those values of p which are positive in the interval in which the diffracted pulse is present. This is certainly less than

$$J_2^* = 2\int_{L-r\cos\theta'}^{\infty} p_0(\tau+r\cos\theta')\,d\tau = 2\int_L^{\infty} p_0(v)\,dv$$

(the diffracted pulse, considered alone, being always negative), and by what has been said above,

$$|J_2| \leqq |J_2^*| \leqq \epsilon.$$

Thus it is found that

$$|2I_0-J| = |2I_0-J_1-J_2| \leqq 2\epsilon,$$

and therefore

$$\lim_{L=\infty} J = 2I_0. \tag{8}$$

When the incident pulse is rectangular and of unit length, assuming its total impulse to be unity, the positive impulse will be equal to 2 when L exceed unity; for smaller values of L, it is given by

$$J = \int_{-r\cos\theta'}^{1-r\cos\theta'} \left\{2-\frac{2}{\pi}\tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}\right\} d\tau = 2-\frac{2}{\pi}\int_r^{1-r\cos\theta'} \tan^{-1}\sqrt{\left(\frac{\tau-r}{L}\right)}\,d\tau$$

or, after a simple integration,

$$J = 2-\frac{2}{\pi}\left\{\tan^{-1}\sqrt{\left(\frac{1-L}{L}\right)}-\sqrt{[L(1-L)]}\right\}.$$

Reflexion by wedges

6. I have discussed some of the features which distinguish a real plane reflector from an infinite reflecting plane by considering the reflexion properties of a semi-infinite screen. Many other 'models' of plane reflectors can be constructed by utilizing Sommerfeld's solution of the problem of the diffraction of plane waves by an infinite wedge, and I propose to give a summary of the results obtained from the consideration of one particular arrangement.

Suppose that the direction of propagation of the incident pulse is fixed, and makes equal angles with the faces of the infinite reflecting wedge. The pattern formed by the wave fronts of the direct, reflected and diffracted pulses is in many respects similar to that of the semi-infinite screen, and is shown in figure 3. There is again, at any moment, a region in which the 'simple theory' holds, and at any point on one of the faces of the wedge there is a 'time lag' between the direct and the diffracted pulse which is equal to the time lag on a semi-infinite screen coinciding with the face of the wedge considered. The diffracted wave will have the same kind of effect as before, it will tend to reduce the 'doubling' of the pressure; and if the point considered is far enough from the edge, the 'simple theory' will hold approximately.

Variations of the angle of the wedge can be taken to represent variations of the angle of incidence. But they change the shape of the reflector at the same time, so that the diffracted wave is modified. A '180° wedge' is an infinite plane (at right angles to the direction of propagation of the incident pulse), there is no diffracted wave and the 'simple theory' holds. As the angle of the wedge is decreased, the diffraction effect—the reduction of the pressure increase due to 'simple' reflexion—increases, and the diffracted wave tends to approximate to the form of the incident wave, but is of course negative; in the limit, when the angle of the wedge is zero, the diffracted pulse just cancels the reflected pulse, and the incident one is propagated undisturbed.

At the edge of the wedge, the total pressure equals the pressure of the incident pulse multiplied by a certain factor depending on the angle of the wedge, which is

$$\frac{2\pi}{\alpha}$$

(where α is the external angle of the wedge).

It can be proved that this is also the ratio of the total impulse, which is the same at all points, to the total impulse of the incident pulse.

In figure 3 pressure-time curves at points on a semi-infinite screen, a 30 and a 90° wedge, are compared for various values of L (the time lag), the incident pulse being rectangular. The areas enclosed between these curves and the time axis, in the same order, are 1·0, 1·091 and 1·333.

It can be shown, by an argument similar to that employed in the case of the semi-infinite screen, that the positive impulse tends to double the total impulse of an incident pulse involving positive excess pressure only, as the distance of the point considered from the edge tends to infinity. At the edge, the positive impulse is in that case equal to the total impulse. For the particular case in which the incident pulse is one in which the pressure first rises discontinuously and then decreases exponentially, the positive impulse has been plotted against distance from the edge in figure 4. The unit of length is here a 'pulse length' defined as the τ interval in which the pressure (of the incident pulse) falls to $1/e$ of its initial value.

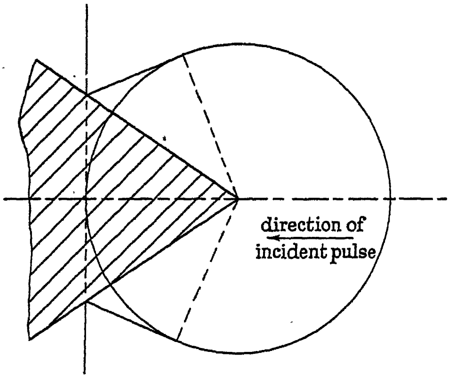


FIGURE 3. The wave fronts in the case of the diffraction of a plane pulse by a symmetrically placed infinite wedge.

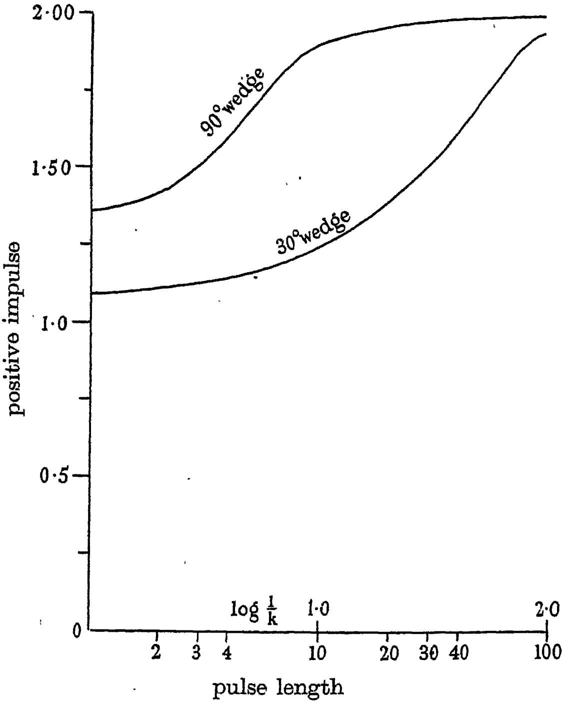


FIGURE 4. 'Positive impulse' plotted against 'pulse length' $1/k$ for the pulse $p_0 \propto e^{-kct}$.

CONCLUSION

7. The main object of the present investigation has been to elucidate the origin of a certain paradox which is a consequence of the simple theory of the reflexion of plane waves by an infinite plane reflector. As has already been pointed out, an infinite plane reflector is a purely mathematical concept whose utility can only be gauged by its success in approximating to conditions which can be physically realized. From this point of view the paradox is a warning that the 'simple theory' is inadequate when the angle of incidence approaches 90°.

By considering the reflexion of plane waves by a semi-infinite screen, and by wedges, it has then been shown that diffraction effects will in that case interfere with the working of the 'simple' theory, which no longer represents the actual state of affairs on such a reflector under very oblique incidence in any way.

It seems reasonable to generalize this argument to a plane reflector of any type. This immediately suggests the question: when is the 'simple theory' a good approximation, i.e. when is a region on the reflecting surface of a plane reflector 'sufficiently' far removed from the edges for it to be treated as part of an infinite reflecting plane? There is clearly no general answer to this question; a comparison of the results obtained for the semi-infinite screen, and for the wedges, suggests that whenever the 'time lag' of the diffracted wave is small (in comparison to some typical dimension of the incident pulse) its effect is appreciable.

In conclusion, I wish to express my thanks to Professor G. I. Taylor for much stimulating discussion and advice in connexion with this paper.

The rate of evaporation of droplets.

Evaporation and diffusion coefficients, and vapour pressures of dibutyl phthalate and butyl stearate

By R. S. Bradley, M. G. Evans and R. W. Whytlaw-Gray, F.R.S.

Department of Chemistry, The University of Leeds.

(*Received* 19 *February* 1945)

The rate of evaporation of drops of dibutyl phthalate and butyl stearate of radius approx. 0·5 mm. has been studied by means of a microbalance over a range of atmospheric pressures down to approx. 0·1 mm. of mercury. Wide departures from Langmuir's evaporation formula were found to occur at these low pressures, but results are in good accordance with the theory of droplet evaporation advanced by Fuchs which hitherto has not been tested experimentally. This experimental verification of Fuch's theory for droplets of medium size evaporating at low pressures shows that the theory can be applied to the evaporation of very small drops at atmospheric pressure. The vapour pressures of the above liquids have been measured by Knudsen's method and the evaporation and diffusion coefficients calculated from the experimental data.

List of symbols used

a = radius of drop.
A = collision area for 'air molecule' and diffusing molecule.
A_1, A_2 = areas of holes in vapour-pressure experiment.
α = evaporation coefficient.
α_{12} = factor to allow for persistence of velocities.
c = concentration of vapour at wall.
c_0 = concentration of vapour at saturation.
c_1 = concentration of vapour at distance Δ from drop surface.
D = diffusion coefficient at air pressure P; $D = Y/P$.
Δ = distance from surface of drop to region of concentration c_1, $= kT/AP$.
H $= \alpha\nu P/Y$.
k = gas constant per molecule.
$k_1 k_2$ = constants in Knudsen's equation (see equations (27), (29)).
L $= \alpha\nu c_0 m_2/\rho$.
L_1 = length of tube in vapour-pressure experiment.
λ_2 = mean free path of a diffusing molecule.
m_1 = mass of 'air molecule'.
m_2 = mass of diffusing molecule.
m = mass of droplet.
M = molecular weight.
ν $= (kT/2\pi m_2)^{\frac{1}{2}}$.
p = pressure of vapour.
p_0 = pressure of vapour at saturation over drop.

p_∞ = pressure of vapour at saturation over a plane surface.
p_1 = $c_1 kT$.
P = pressure of air.
q = ds/dt.
q_0 = ds/dt for portion of graph of ds/dt versus $1/P$ which is linear.
r = radial distance from droplet centre.
r_0 = distance of charcoal from droplet centre.
R = gas const. per g.mol.
$R_1 R_2$ = radii of openings in vapour-pressure experiment.
ρ = density of evaporating liquid.
S_{12} = sum of radii of air molecule and diffusing molecule.
s = surface of drop.
s_0 = surface of drop at zero time.
σ = surface tension of liquid.
T = absolute temperature.
t = time.
u = velocity with which molecules leave the droplet surface.
W = $av\alpha/Y$.
$W_1 W_2$ = weights of liquid lost by evaporation in vapour-pressure experiment.
x = constant to allow for tube resistance in vapour-pressure experiment.
Y = DP.

I. INTRODUCTION

It is well known that small spheres of liquid evaporate in air in such a manner that the rate of change of surface with time is constant, a result which, as Langmuir (1918) showed, follows directly from Stephan's theory of diffusion. The expression

$$-ds/dt = 8\pi D m_2 c_0/\rho \tag{1}$$

(where s = surface, t = time, D = diffusion coefficient, c_0 = concentration in the saturated vapour, m_2 = mass of a diffusing molecule, and ρ = density of the liquid) has been verified experimentally for droplets varying in radius from a few millimetres to about 0·1 mm., and tested for a number of liquids of varying vapour pressure. It has also been tested, though less rigorously, on much smaller droplets in the size range 2×10^{-4} to 5×10^{-5} cm. by observing their rates of fall through air at atmospheric pressure. For these small spheres high-boiling organic substances were used, so that evaporation took place at a sufficiently low rate for measurement, and the interesting observation was made that the graph of s against t, instead of being linear, curved slightly away from the time axis, the curvature being more marked the smaller the particle. It was clear that evaporation of these small droplets showed a progressive diminution in rate (expressed as $-ds/dt$) as smaller dimensions were reached. At the time this behaviour was attributed to the presence of non-volatile impurities, but in the light of recent work it seems likely that it really represents the normal course of evaporation of spheres of these sizes. Unfortunately, on account of experimental difficulties, the experiments were not repeated and a study of still smaller droplets which should have shown a greater deviation proved still more difficult on account of increasing Brownian movement (Whytlaw-Gray & Patterson 1932, p. 169).

In addition, experiments have been made in these laboratories on the evaporation in air, at various pressures down to about 1 cm. of mercury, of droplets of about 1 mm. radius, and it was observed that whilst at any one pressure ds/dt remained constant its magnitude varied inversely as the pressure (P), a result in conformity with the diffusion theory, since D is proportional to $1/P$. It is, however, clear that ds/dt cannot increase indefinitely with falling pressure, for evaporation can never exceed the vacuum rate; ds/dt must therefore approach asymptotically the rate of evaporation in a perfect vacuum as the air pressure becomes smaller and smaller.

Similarly, for droplets evaporating in air at atmospheric pressure, there must be a limiting size below which the rate of evaporation ceases to follow the simple expression, for, if it were not so, the number of molecules escaping per unit area from a very small droplet would exceed the number escaping into a vacuum, which is obviously impossible. This follows from Langmuir's formula

$$-\frac{dm}{dt} = 4\pi a D c_0 m_2, \tag{2}$$

where a is the radius and m the mass of the drop. Hence the loss per unit area per second is Dc_0m_2/a, which approaches ∞ as a approaches zero. It is hence clear that Langmuir's simple expression cannot be valid either for the evaporation of very small droplets at atmospheric pressure or for larger droplets at very low pressures. It is, however, reliable for droplets down to 10^{-3} cm. radius in normal air, and the deviations even at 10^{-4} cm. radius are small. At lower pressures for the larger drops the departure of ds/dt from constancy is inconsiderable down to about 1 cm. of mercury.

The theoretical aspect of droplet evaporation has been discussed by Fuchs (1934). Fuchs regards the diffusion process as not starting directly at the surface of the evaporating sphere of radius a but from the surface of an enveloping sphere of radius $a+\Delta$, where Δ is of the order of the mean free path of the air molecules. Very few air molecules will then be present in the spherical shell of thickness Δ, which is substantially a vacuous space. The effect of this space will be negligibly small for larger droplets, but when a approaches Δ in magnitude the rate of evaporation expressed as $-ds/dt$ will be decreased. Starting from this conception, Fuchs developed equations for the rate of evaporation of very small spheres but was unable to check them on account of the paucity of experimental data.

The present report presents a more complete treatment of the theory of evaporating droplets and describes experimental work carried out in order to throw light on the course of evaporation of very small liquid spheres. Since the direct study of small droplets is impracticable, we have had recourse to observations on comparatively large drops (*ca.* 0·5 mm. radius) at low pressures. Our treatment of the theory shows that information so gained is directly applicable to the calculation of the evaporation of very small droplets at atmospheric pressure.

II. THEORETICAL

(*a*) *The general evaporation equation: drop in an infinite enclosure*

Following the treatment of Fuchs we suppose that the concentration of vapour at a distance $a+\Delta$ from the centre of the drop is c_1 molecules per c.c., a value less than the equilibrium concentration c_0, as will be seen below; Δ is, of course, an equivalent distance, since actually there must be a concentration gradient within the spherical shell of radii a and $a+\Delta$. The drop is imagined to evaporate in a gas at pressure P into an infinite space.

If the evaporation or accommodation coefficient is α, i.e. a fraction α of the molecules which hit the surface condense, then the rate of evaporation into a vacuum is $4\pi a^2 \nu \alpha c_0$ molecules per second, where $\nu = (kT/2\pi m_2)^{\frac{1}{2}}$, k being the gas constant per molecule, T the absolute temperature, and m_2 the mass of the evaporating molecule. Hence the rate at which molecules arrive at the shell surface, distant $a+\Delta$ from the droplet centre, is $4\pi a^2 \nu \alpha (c_0 - c_1)$. This may be equated to the rate at which molecules leave by diffusion, in the steady state, or

$$4\pi a^2 \nu \alpha (c_0 - c_1) = 4\pi (a+\Delta) D c_1,$$

where D is the diffusion coefficient at the pressure P. Hence

$$c_1[(a+\Delta) D + a^2 \nu \alpha] = c_0 a^2 \nu \alpha,$$

giving an evaporation rate

$$-\frac{dm}{dt} = 4\pi (a+\Delta) \frac{D c_0 a^2 \nu \alpha m_2}{(a+\Delta) D + a^2 \nu \alpha},$$

where m is the mass of the drop, or

$$-\frac{dm}{dt} = \frac{4\pi a D c_0 m_2}{D/(a\nu\alpha) + a/(a+\Delta)}. \tag{3}$$

Fuchs left the evaluation of Δ in an indefinite form, but a sufficiently accurate result is easily derived. Let the collision area for the impact of evaporating and air molecules be A, i.e. A is a mean value for the constituents of air. Let the velocity with which the diffusing molecules leave the droplet be u, in general greater than the mean thermal value. The number of collisions is $u(AP/kT)$ per second, the time between collisions kT/AuP, and the distance between collisions

$$(kT/AuP)\, u = kT/AP = \Delta,$$

i.e. a molecule after leaving the droplet travels a distance Δ before a collision occurs.

Since λ_2, the mean free path of evaporating molecules, is given by

$$\lambda_2 = \left(\frac{m_1}{m_1+m_2}\right)^{\frac{1}{2}} \frac{kT}{\pi P S_{12}^2},$$

where m_1 is the mean mass of an air molecule, S_{12} the sum of the radii of evaporating and 'air molecules', and the concentration of evaporating molecules is neglected, then

$$\frac{\Delta}{\lambda_2} = \left(\frac{m_1+m_2}{m_1}\right)^{\frac{1}{2}} \frac{\pi S_{12}^2}{A}$$

$$= \left(\frac{m_1+m_2}{m_1}\right)^{\frac{1}{2}}.$$

It should be noted that, unlike Langmuir's simple expression, equation (3) gives the correct rate of evaporation into a vacuum for drops of radius of the order 1 mm. For then Δ is large compared with a and $a/(a+\Delta)$ vanishes. Hence

$$-dm/dt = 4\pi a^2 \nu \alpha c_0 m_2,$$

i.e. the correct value for $a = 1$ mm. and $P = 0$.

On the other hand, Langmuir's expression is reproduced under the conditions for which it holds, i.e. a of the order 1 mm. and $P = 1$ atm. Then $a/(a+\Delta)$ becomes nearly unity and $D/(a\nu\alpha)$ is negligible compared with unity, if a is mm., D approx. 0·1, α approx. unity and $T = 20°$ C, taking a molecular weight of the order 100.

It is clear, however, that $D/(a\nu\alpha)$ will not be small compared with unity if either a or α is small. Considering only the variation in a at constant P and constant D, equation (3) gives for the rate of evaporation *per unit area*

$$-\frac{dm}{dt}\,(\text{per unit area}) = \frac{Dc_0 m_2}{\dfrac{D}{\nu\alpha} + \dfrac{a^2}{a+\Delta}}$$

$$= \nu\alpha c_0 m_2 \quad \text{if } a \text{ is very small.}$$

This is the vacuum rate, and it follows, therefore, that the rate of evaporation of very small drops at atmospheric pressure is, *per unit area*, nearly equal to the rate of evaporation of unit area into a vacuum. This is clearly a remarkable and important result. As already mentioned, Langmuir's expression gives for the rate of evaporation per unit area $-dm/dt$ (per unit area) $= Dc_0 m_2/a$, i.e. infinite values as $a \to 0$. Fuch's theory therefore gives evaporation rates per unit area at gas pressure P *less* than those given by Langmuir, but *greater* than those calculated for a plane sq.cm. evaporating into gas at pressure P.

Clearly, equally great departures from Langmuir's result would be expected if the evaporation coefficient were very small, for then $D/\nu\alpha a$ would be large compared with $a^2/(a+\Delta)$ even when $a = 1$ mm. and $P = 1$ atm. This is not likely to occur, however, with pure liquids.

These results are summarized in table 1, which contrasts the results of Fuchs

TABLE 1

equation due to	decreasing radius (a) at constant pressure		
	$-\frac{dm}{dt}$	$-\frac{dm}{dt}\frac{1}{4\pi a^2}$ (per unit area)	$-\frac{ds}{dt}$
Langmuir	$4\pi a D c_0 m_2 \to 0$ as $a \to 0$	$\frac{D c_0 m_2}{a} \to \infty$ as $a \to 0$	$\frac{8\pi D m_2 c_0}{\rho}$ constant as $a \to 0$.
Fuchs	$\frac{4\pi a D c_0 m_2}{\frac{D}{a v \alpha}+\frac{a}{a+\Delta}} \to 0$ as $a \to 0$	$\frac{D c_0 m_2}{\frac{D}{v \alpha}+\frac{a^2}{a+\Delta}} \to v c_0 \alpha m_2$ as $a \to 0$ (vacuum rate)	$\frac{8\pi D m_2 c_0}{\rho}\left\{\frac{1}{\frac{D}{a v \alpha}+\frac{a}{a+\Delta}}\right\} \to 0$ as $a \to 0$
	decreasing pressure (P) at constant radius		
	$-\frac{dm}{dt}$	$-\frac{dm}{dt}\frac{1}{4\pi a^2}$	$-\frac{ds}{dt}$
Langmuir	$\frac{4\pi a Y c_0 m_2}{P} \to \infty$ as $P \to 0$	$\frac{Y c_0 m_2}{a P} \to \infty$ as $P \to 0$	$\frac{8\pi Y c_0 m_2}{P \rho} \to \infty$ as $P \to 0$
Fuchs	$\frac{4\pi a c_0 m_2}{\frac{1}{a v \alpha}+\frac{P a}{Y(a+\Delta)}} \to 4\pi a^2 v c_0 \alpha m_2$ as $P \to 0$ (vacuum rate)	$\frac{c_0 m_2}{\frac{1}{v \alpha}+\frac{P a^2}{Y(a+\Delta)}} \to \alpha v c_0 m_2$ as $P \to 0$ (vacuum rate)	$\frac{8\pi Y m_2 c_0}{P \rho}\left\{\frac{1}{\frac{Y}{P a v \alpha}+\frac{a}{a+\Delta}}\right\} \to \frac{8\pi a v \alpha c_0 m_2}{\rho}$ as $P \to 0$ (vacuum rate)

and Langmuir under various conditions. Fuch's equation may be transformed to give the rate of change of surface with time in the form

$$-\frac{ds}{dt}=\frac{8\pi Dm_2c_0}{\rho}\left(\frac{1}{\dfrac{D}{av\alpha}+\dfrac{a}{a+\Delta}}\right). \tag{4}$$

The diffusion coefficient D is inversely proportional to the pressure, i.e. $D = Y/P$, where Y is a constant.

(b) *The theory of evaporation in a finite vessel with absorbent walls*

It now remains to study the evaporation of droplets under the conditions used in the experiment, in which a droplet is suspended inside an enclosure, the walls of which consist of an absorbent material such as charcoal, giving effectively zero vapour pressure at the walls. It is considered that the equilibrium pressure over the charcoal is roughly the same as that over the evaporating substance in the solid state at some temperature below the melting point, and owing to the large temperature coefficient of vapour pressure of the liquids used the vapour pressure at the walls is effectively zero. First let us derive Langmuir's result under experimental conditions, and then consider modifications.

(i) From simple diffusion theory, neglecting Fuchs's correction for the moment,

$$\frac{dm}{dt}=4\pi r^2Dm_2\frac{dc}{dr},$$

where c is the concentration in molecules per c.c. at a distance r from the drop centre.

If the last equation is written in the form $\left(\frac{dm}{dt}\right)\frac{dr}{4\pi r^2}=Dm_2dc$ and integrated,

$$-\frac{dm}{dt}\left(\frac{1}{a}-\frac{1}{r_0}\right)=4\pi Dm_2c_0, \tag{5}$$

where the concentration is zero at a distance r_0 from the drop centre, and a is the droplet radius; it is assumed that the slight variation in total pressure from the droplet surface to the wall does not affect the diffusion coefficient D, which is therefore regarded as a constant during the integration. It is also assumed that the mass of vapour passing through every spherical surface in unit time is independent of the radius of the surface, i.e. that the steady state is reached. In experimental practice the enclosure quickly fills with vapour at a graded concentration, and the steady state is achieved in a short time. Moreover, put $p_0 = c_0kT$, since p_0 is usually small.

If $r_0 = \infty$ equation (5) may be written

$$-4\pi a\rho\frac{da}{dt}=\frac{4\pi Dm_2p_0}{kT},\quad \text{i.e.}\quad -\frac{ds}{dt}=\frac{8\pi}{\rho}\frac{Dm_2}{kT}p_0, \tag{6}$$

which is Langmuir's result. It follows, therefore, that Langmuir's formula is not strictly true with a finite enclosure. The direct integration of equation (5) gives

$$\frac{a^2}{2}-\frac{a^3}{3r_0}=-\frac{Dm_2p_0t}{kT\rho}+\frac{a_0^2}{2}-\frac{a_0^3}{3r_0}, \tag{7}$$

where $a = a_0$ when $t = 0$. Hence it is not quite true to say that the rate of change of surface with time is constant. In many experiments, however, a changes only slightly with time during the run, and thus

$$-\frac{1}{a}\left(\frac{dm}{dt}\right) = \frac{r_0}{r_0-a}\frac{4\pi D m_2 p_0}{kT}, \tag{8}$$

or

$$s = \frac{-8\pi D m_2 p_0 t}{\rho kT\left(1-\frac{a}{r_0}\right)} + s_0. \tag{9}$$

This equation is identical with that due to Langmuir, except for the introduction of a factor $\left(\frac{1}{1-\frac{a}{r_0}}\right)$. When a is increased to infinity, and $r_0 - a$ kept constant, then the correct expression for a plane surface for dm/dt per unit area is reproduced.

(ii) Now consider how these equations which apply to experimental conditions are modified by the conceptions of Fuchs.

When Fuchs's correction is introduced then, instead of equation (5), it is seen that

$$-\frac{dm}{dt}\left(\frac{1}{a+\Delta}-\frac{1}{r_0}\right) = 4\pi D c_1 m_2,$$

where $p_1 = c_1 kT$, and $P \gg p_1$; c_1 is the concentration of vapour at the surface of radius $a+\Delta$. Hence the rate of loss of mass by diffusion is

$$-4\pi D c_1 m_2 \frac{(a+\Delta)\,r_0}{(r_0-a-\Delta)} \text{ g. per sec.}$$

But this must be balanced by the rate of arrival of molecules from the surface of the droplet to the region of concentration c_1 or

$$4\pi a^2 \nu\alpha(c_0-c_1) = \frac{4\pi D c_1 (a+\Delta)\,r_0}{(r_0-a-\Delta)}.$$

Hence

$$c_1 = \frac{a^2\nu\alpha c_0}{\frac{(a+\Delta)\,D r_0}{(r_0-a-\Delta)} + a^2\nu\alpha},$$

and

$$-\frac{dm}{dt} = \frac{4\pi a D c_0 m_2 r_0/(r_0-a-\Delta)}{\frac{D r_0}{(r_0-a-\Delta)\,a\nu\alpha} + \frac{a}{a+\Delta}}$$

$$= \frac{4\pi a Y c_0 m_2}{P\left[Y/(Pa\nu\alpha) + \left\{1-\frac{(a+\Delta)}{r_0}\right\}\frac{a}{a+kT/(AP)}\right]}$$

$$= \frac{4\pi a Y c_0 m_2}{P\left[Y/(Pa\nu\alpha) + a\left(\frac{AP}{aAP+kT}-\frac{1}{r_0}\right)\right]}. \tag{10}$$

Since $\Delta = kT/(AP)$.

Hence $$-a(da/dt) = \frac{Yc_0m_2}{\rho P\left[Y/(Pav\alpha)+a\left(\dfrac{AP}{aAP+kT}-\dfrac{1}{r_0}\right)\right]}$$

or $$a+H\{\tfrac{1}{2}(a+\Delta)^2-2\Delta(a+\Delta)+\Delta^2\log(a+\Delta)\}-Ha^3/3r_0 = -Lt+\text{constant},$$

or $$a+H\left\{\frac{a^2}{2}-a\Delta+\Delta^2\log(a+\Delta)\right\}-\frac{Ha^3}{3r_0} = -Lt+\text{constant}, \tag{11}$$

where $$H = \alpha vP/Y = \alpha v/D \quad\text{and}\quad L = \alpha vc_0m_2/\rho.$$

It should be noted that if the drop evaporates into a region of constant vapour pressure c, then

$$-dm/dt = 4\pi(a+\Delta)\,Dm_2\left[\frac{c_0a^2v\alpha}{(a+\Delta)\,D+a^2v\alpha}-c\right].$$

Hence no simple allowance can be made for the influence of the vapour pressure corresponding to concentration c, as can be done for Langmuir's expression, viz.

$$-\frac{dm}{dt} = 4\pi aDm_2(c_0-c).$$

A graphical integration is necessary if the change of radius with time is required for $c>0$.

Although equation (11) is the correct form of integration, and should replace the simple equation (9), in much of the work described below the radius of the drop changes only slightly during the run, owing to the sensitive nature of the balance used. Hence equation (10) may be integrated as if the correcting denominator were constant during the run, giving

$$s-s_0 = \frac{-8\pi Ym_2c_0t}{\rho P\left[Y/(Pav\alpha)+a\left(\dfrac{AP}{aAP+kT}-\dfrac{1}{r_0}\right)\right]}. \tag{12}$$

Hence, under the conditions for which a changes very little during a run, we again have the familiar law, s a linear function of time, but the proportionality constant varies with pressure in a manner very different from that given by equation (9). The latter gives ds/dt proportional to $1/P$, since $D = Y/P$, i.e. indefinite increase of ds/dt as P decreases, whereas equation (12) gives an asymptotic approach of ds/dt to the vacuum rate. Equation (12) is fully substantiated by the experimental work below. If, however, a varies considerably during the run then equation (11) must be applied. With drops such as were used for runs in which the radius changed little during the run only the first two terms are significant, but this would not be the case with small drops of radius 0·01 cm. at low pressures (e.g. 0·3 mm. of mercury), as is shown experimentally later.

One may note that at very low pressures of air, or at vapour pressures comparable with air pressures, the above analysis will not apply, since D will not be a constant between the droplet surface and the wall, owing to the change of D with pressure. No simple solution for c_1 can be given. However, one can extrapolate equations

(11) and (12) to $p = 0$ with confidence, since the value of ds/dt is very insensitive to changes of P when P is very low, and the difference between the value of ds/dt calculated from equation (12) and the true result must be very small.

It is assumed that p_0 is independent of P, which is not quite true, since $\frac{dp_0}{dP} = \frac{v_{\text{liquid}}}{v_{\text{vapour}}}$, i.e. dp_0 = of the order of 1 % for $dP = 1$ atm. for dibutyl phthalate. Another point of interest in this connexion is the formation of adsorption complexes in the gaseous phase, owing to the solvent effect of the atmosphere on the vapour. Owing to the large van der Waals' attractive force of a molecule such as dibutyl phthalate, it is quite possible that in air a cluster of molecules is formed which may be regarded as air molecules adsorbed on the dibutyl phthalate molecule, so that m_2 is not strictly constant as the air pressure changes.

Equation (12) may be tested by plotting ds/dt as a function of $1/P$ (a changes only slightly during the run). For large values of P the plot is a straight line, corresponding to Langmuir's case, but as P decreases the correcting term becomes more and more important. If $ds/dt = q$ for some value of P, and

$$\frac{-8\pi Y m_2 c_0 t}{\rho P} = q_0, \tag{13}$$

q_0 is readily calculated from the linear portion of the graph, and

$$\frac{q_0}{q} = \frac{Y}{Pav\alpha} + a\left(\frac{AP}{aAP+kT} - \frac{1}{r_0}\right) \tag{14}$$

$$= \frac{1}{PW} + \frac{1}{1+kT/(aAP)} - \frac{a}{r_0}, \tag{15}$$

where $W = av\alpha/Y$.

Now A can be roughly computed from the dimensions of the molecule, and the term $\frac{1}{1+kT/(aAP)}$ is thus known. Under the conditions of the experiment the calculation is not very sensitive to changes in A. Then if $\frac{q_0}{q} - \frac{1}{1+kT/(aAP)}$ be plotted as a function of $1/P$ a straight line should be obtained, from which W may be calculated. If the vapour pressure is known, Y may be calculated from ds/dt for large values of P, and thence α may be computed, since $W = av\alpha/Y$.

Alternatively, Y may be calculated from diffusion theory. Unfortunately, the latter is not very exact, unless the complicated approach of Chapman and Enskog is adopted. The equation of Meyer, modified to allow for persistence of velocities, seems to be most suitable for calculation. According to this

$$D = \frac{2}{3(1+\alpha_{12})}\frac{kT}{\pi P S_{12}^2}\left[\frac{2kT}{\pi}\frac{(m_1+m_2)}{(m_1 m_2)}\right]^{\frac{1}{2}}, \tag{16}$$

$$Y = \frac{2}{3(1+\alpha_{12})}\frac{kT}{\pi S_{12}^2}\left[\frac{2kT}{\pi}\frac{(m_1+m_2)}{(m_1 m_2)}\right]^{\frac{1}{2}}, \tag{17}$$

where α_{12} allows for the persistence of velocities, and is a small term of approximate value $m_1/(m_1+m_2)$ under the conditions of the experiment. Maxwell's exact solution for the case of repulsion between molecules with an inverse fifth power law gives the same dependence on P, but will not be used since the force constants are unknown and since the Maxwell law is in any case unreal. It may be noted that according to equation (16) D is proportional to $T^{\frac{3}{2}}$ at constant P, whereas experimentally the index of T is greater than 1·5, for water 1·75.

It is of some interest to consider modifications to the above due to the influence of surface tension on vapour pressure. If σ is the surface tension and p_∞ the saturation vapour pressure over a plane surface

$$\log\frac{p_0}{p_\infty}=\frac{2\pi m_2\sigma}{kTa\rho}$$

or

$$p_0=p_\infty e^{\frac{2\pi m_2\sigma}{kTa\rho}}. \tag{18}$$

Equation (10) becomes, after putting $p_0=c_0kT$,

$$-\frac{dm}{dt}=\frac{4\pi aYp_\infty m_2 e^{\frac{2\pi m_2\sigma}{kTa\rho}}}{kTP\left[Y/(Pav\alpha)+a\left(\frac{AP}{aAP+kT}-\frac{1}{r_0}\right)\right]}. \tag{19}$$

Hence

$$da\left[Y/(\nu\alpha)+a^2P\left(\frac{AP}{aAP+kT}-\frac{1}{r_0}\right)\right]e^{\frac{-2\pi m_2\sigma}{kTa\rho}}=\frac{-Yp_\infty m_2 dt}{kT\rho}. \tag{20}$$

Provided that $p_0/p_\infty-1$ is $\ll 1$ one can expand the exponential and integrate, giving, if the term in r_0 be omitted,

$$a+H\{\tfrac{1}{2}a^2-a\Delta+\Delta^2\log(a+\Delta)\}\frac{-2\pi m_2\sigma}{kT\rho}[\log a+H\{a-\Delta\log(a+\Delta)\}]$$

$$=\frac{-\alpha\nu p_\infty m_2 t}{kT\rho}+\text{constant}. \tag{21}$$

As the surface tension effect does not become appreciable before the radius is very small, the droplet has practically vanished before this pressure condition is violated.

III. EXPERIMENTAL. RATES OF EVAPORATION

In order to obtain great sensitivity a quartz microbalance A (figure 1) of the type described by Whytlaw-Gray was used. To one end of the balance was attached a long quartz fibre, terminating in a hook from which was suspended the quartz rod carrying the drop. The balance was counterpoised with quartz weights; detail of the balance is given inset in figure 1. The movement of the pointer B relative to the reference pointer C could be followed by observation through the plate glass window D by means of a travelling microscope reading to 0·001 mm. The window D was ground to fit the pyrex balance case. Surrounding the drop was a copper gauze cage of radius 1·2 cm., and charcoal was packed in the annular space between the cage and the vessel *E*. The balance case was mounted on a massive block of metal, and covered with a metal hood, to prevent temperature fluctuations, and the whole apparatus was placed in a thermostated room at 20°. *F* was a sintered glass filter to filter particles from the air, *G*, *G* were P_2O_5 tubes, and *H* a manometer which was read to 0·01 mm. Low pressures were read on the McLeod gauge. The balance was cleaned by $HNO_3 + HCl$, the glass apparatus by $KMnO_4$ + concentrated H_2SO_4.

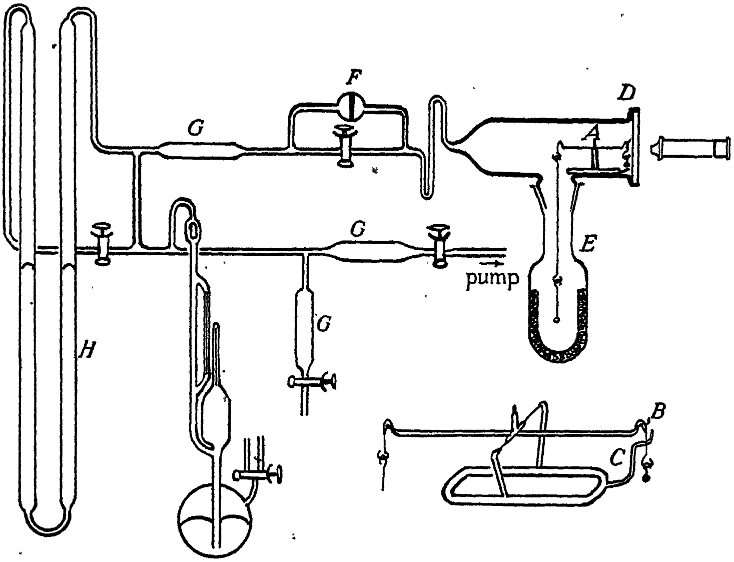


FIGURE 1

The end of the quartz rod carrying the drop was drawn out to a very thin fibre, and drops could be made to adhere to the end of this fibre by first forming them at the end of another fibre, touching the fibre points, and then moving them apart. In this way the fibre was uncontaminated with liquid, except at the end. At first the final counterpoise was effected by means of very small quartz weights, but later it was found easier to adjust the weight of the drop. If a clean quartz fibre be plunged through the drop, at right angles to the fibre carrying the drop, and withdrawn, the weight of the drop is slightly reduced. Conversely, the drop may be made to pick up beads of liquid if a fibre carrying such beads is plunged through the drop. Drops were weighed on a micro-analytical balance weighing to 10^{-6} g., and the quartz balance was calibrated using drops of dibutyl phthalate of different sizes, the weights of which had been determined on the micro-analytical balance. The deflexion versus microscope reading was linear and 1 mm. of the microscope corresponded to $2·85 \times 10^{-5}$ g., i.e. a change of weight of $2·85 \times 10^{-8}$ g. could be observed. Apiezon grease previously dehydrated was used throughout the apparatus. All air admitted to the apparatus was dried by P_2O_5.

Runs were first made with benzophenone, but the rate of evaporation was found to be inconveniently fast at low pressures. Dibutyl phthalate evaporated much more slowly, and using the sensitive balance it was possible to complete the series of runs over a pressure range of 10–0·1 mm. in a day. With butyl stearate several days were necessary to complete the readings. A repetition of experiments gave good agreement, and no change was observed if the liquids were previously degassed hot *in vacuo*. It is believed that in view of the extremely low vapour pressure of these liquids no correction is necessary for surface cooling, especially as the two liquids behaved similarly, although the rate for butyl stearate was some fifteen times less than that for dibutyl phthalate. The general trend of results indicates, also, that surface diffusion over the quartz was negligible (cf. especially table 3).

After the drop has been hung on the quartz balance and the apparatus had been assembled, the whole was pumped down for half an hour to an hour to remove any water vapour which is inevitably introduced when the apparatus is opened to the air. Dry air was then admitted to the required pressure, and after about 30 min. deflexion-time readings taken, and the whole repeated at lower pressures. At low pressures some time elapsed before the charcoal attained approximate adsorption equilibrium with respect to air, and readings could be taken; pressure readings were taken throughout the run, to test the constancy of the pressure. Small percentage variations in pressure occurred, although it was occasionally necessary to pump off air which had been desorbed from the charcoal to maintain the constancy. In this low-pressure region large changes in pressure make small changes in rate of evaporation, and hence the pressure was sufficiently static to give negligible variations in rate due to pressure changes.

Runs using dibutyl phthalate

Comparatively large drops were first used, of radii 0·5 mm., and the variation in radius over the whole of a series of runs at different pressures was approximately 5 % only. During one run at constant pressure the variation in a was 1 % or less, so that equation (12) may be applied with confidence. All runs of this type gave linear plots of surface versus time at constant pressure, even at low pressures, as is seen from figure 2, for a pressure of 0·246 mm.; it will be seen that the change in s over the whole run is small and hence the conditions of equation (12) apply. A summary is given in table 2. This constancy of ds/dt indicates that in all cases a steady state had been attained.

In figure 3 these values of ds/dt are plotted as a function of $1/P$. As will be seen, at first the relation is linear, in accordance with equation (9), i.e. Langmuir's formula, but as P decreases a marked bending away from the initial line occurs. When $q_0/q - \dfrac{1}{1+kT/(aAP)}$ is plotted against $1/P$ a good straight line is obtained (figure 4), in accordance with equation (15), whence the constants of the curve on figure 3 may be calculated. The line goes nearly through the origin, and hence a/r_0 will be

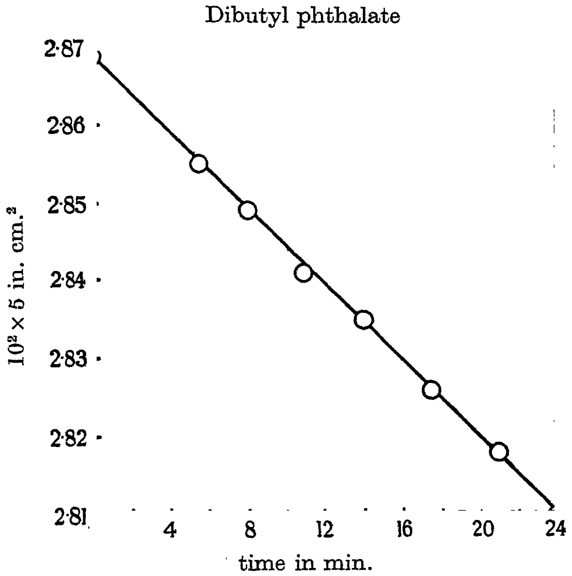


FIGURE 2. Air pressure 0·246 mm.

TABLE 2. DIBUTYL PHTHALATE AT 20°

mean a during run mm.	P mm. of Hg	$-10^5 \times ds/dt = q$ cm.^{2} per min. observed	q_0/q	$-10^5 \times ds/dt$ calc. from equation (22) cm.^{2} per min.
0·5150	8·26	0·192	1·001	0·183
0·5210	3·68	0·409	1·053	0·390
0·5117	3·11	0·419	1·214	0·453
0·5155	1·035	1·115	1·372	1·107
0·5081	0·823	1·317	1·460	1·310
0·4992	0·306	2·319	2·230	2·281
0·5052	0·246	2·435	2·642	2·492
0·4850	0·162	2·944	3·320	2·858
0·4732	0·135	2·899	4·044	2·999
0·4951	0·103	2·988	5·142	3·166

neglected in equation (15). From the dimensions of the molecule and of a mean 'air molecule' the collision radius was taken as 6·28 A. The curved line in figure 3 is calculated from these constants, the equation to the line being

$$-10^5 \times ds/dt = \frac{1{\cdot}583}{p} \frac{1}{\left[\dfrac{0{\cdot}43}{P} + \dfrac{1}{1 + 2{\cdot}444 \times 10^{-3}/(aP)}\right]}, \tag{22}$$

where P is in mm. of Hg, s in cm.^{2}, t in min. The points are experimental and lie very near the line. Thus equation (12) is substantiated. The choice of A is somewhat arbitrary, but fortunately equation (12) in this region of pressure range is rather insensitive to change in A. The mean value of a was approximately 0·05 cm., so that equation (22) may be written in the more general form, applicable to any radius,

$$-10^5 \times ds/dt = \frac{1{\cdot}583}{P} \frac{1}{\left[\dfrac{2{\cdot}15 \times 10^{-2}}{aP} + \dfrac{1}{1 + 2{\cdot}444 \times 10^{-3}/(aP)}\right]}, \tag{23}$$

where P is in mm. of Hg, s in cm.^{2}, a in cm., and t in min.

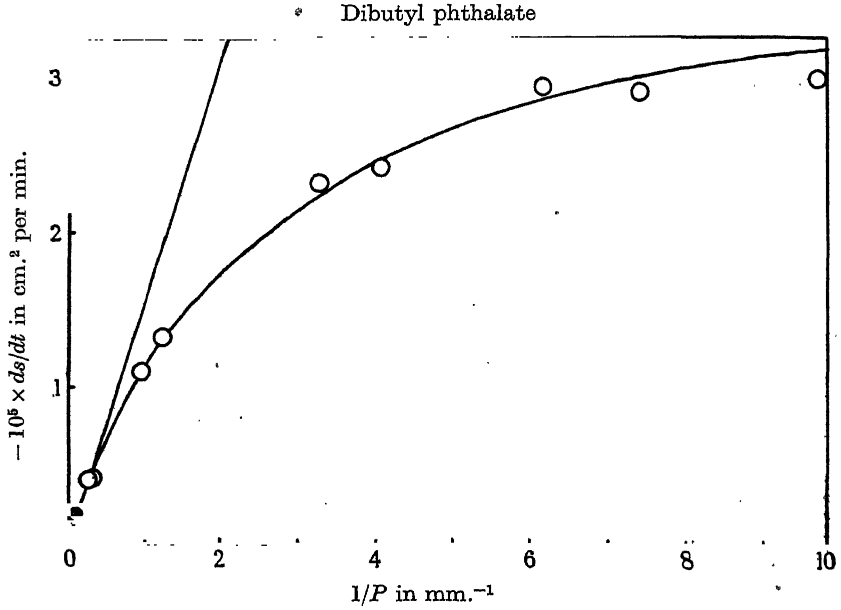


FIGURE 3

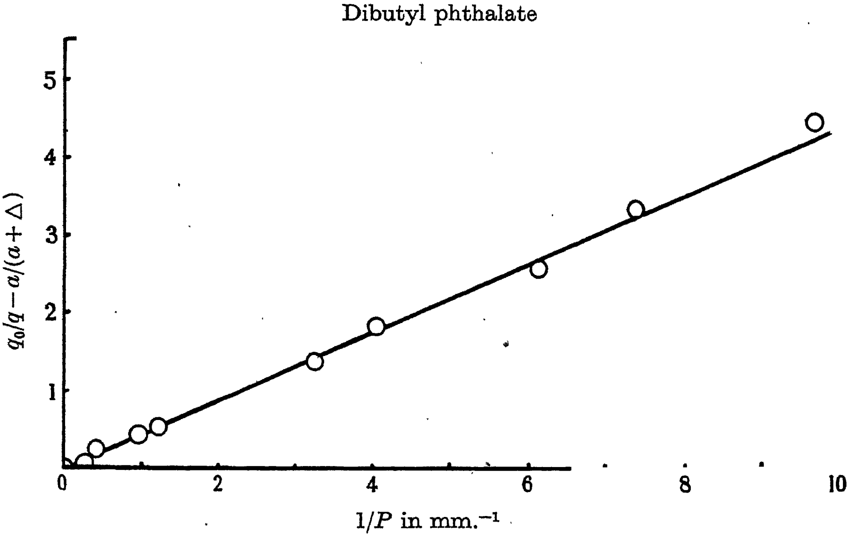


FIGURE 4

The values of q_0/q show the magnitude of the departures from Langmuir's theory. Thus at 0·103 mm. pressure the rate of evaporation is approximately one-fifth of that demanded by Langmuir's theory.

Equation (23) may be used to test the theory of a run in which the radius changes appreciably, when ds/dt should not be constant, but should vary throughout the run. This was found to be true, results for a typical run on a small drop which is allowed to evaporate almost completely being given in table 3. The pressure was 0·309 mm.

TABLE 3

10^7 × wt. of drop g.	time in min.	radius in mm.	10^2 × (radius)^2 mm.^{2}	$10^3\phi(a)$ (equation (24))
584·2	0	0·2370	5·617	1·527
531·7	17	0·2295	5·266	1·462
480·1	34	0·2220	4·928	1·396
417·7	55	0·2119	4·491	1·311
355·3	77	0·2008	4·033	1·216
302·8	97	0·1903	3·623	1·132
241·0	125	0·1764	3·112	1·016
167·2	166	0·1562	2·440	0·856
135·8	184	0·1457	2·123	0·768
106·0	207	0·1342	1·801	0·677
74·9	234	0·1195	1·427	0·571
42·2	271	0·0987	0·974	0·406
24·8	299	0·0826	0·682	0·286
16·1	315	0·0716	0·513	0·205

Figure 5 shows the lack of constancy in $ds/dt = 4\pi d(a^2)/dt$; as is seen from figure 6 da/dt is more constant, but deviates at low values of a. In figure 7 is plotted the

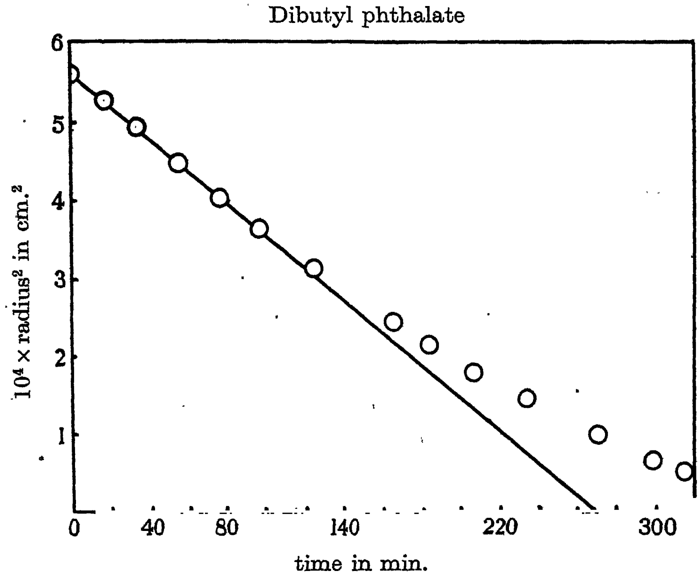


FIGURE 5

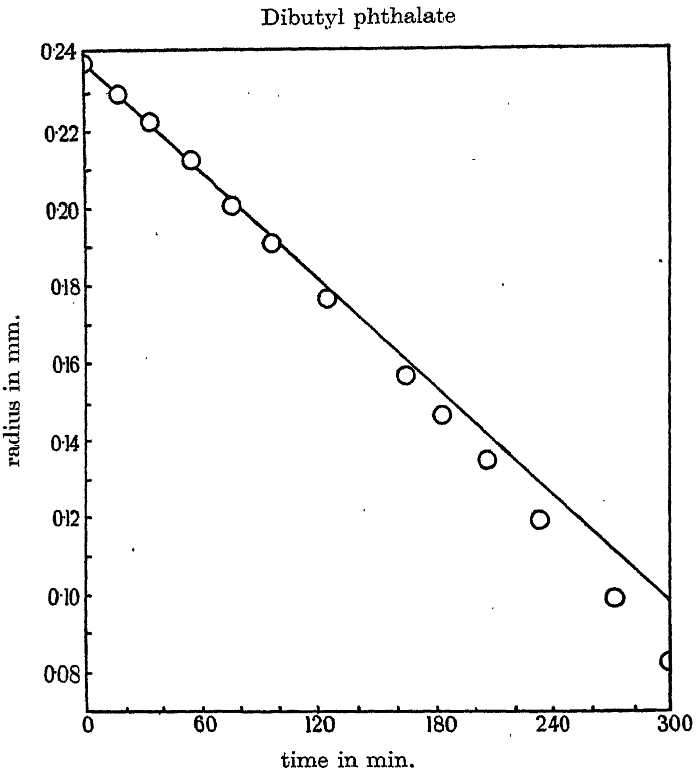


Figure 6

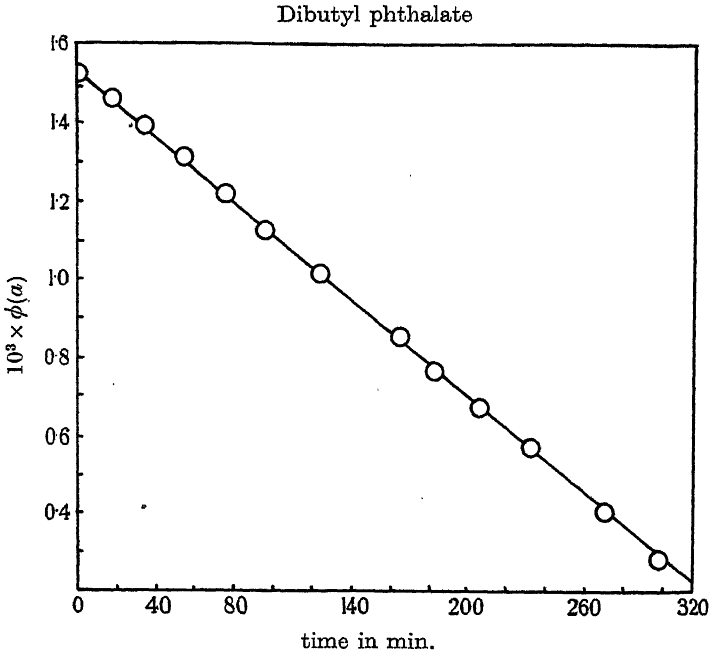


Figure 7

integrated expression from equation (11), expressed in the form, using the constants from equation (23),

$$6{\cdot}167\times10^{-2}a+0{\cdot}5a^2+1{\cdot}436\times10^{-4}\log_{10}(a+7{\cdot}895\times10^{-3})=\phi(a)$$
$$=\text{const.}\,t+\text{const.}\quad(24)$$

It will be seen that $d\phi(a)/dt$ is indeed a constant.

In one respect there is a discrepancy between the data of tables 2 and 3. From equation (23) it may be calculated that $d\phi(a)/dt=-2{\cdot}038\times10^{-6}$, t being measured in minutes, whereas the slope of the line in figure 7 gives $d\phi(a)/dt=-4{\cdot}13\times10^{-6}$. Likewise there is a discrepancy of the same order when the slope $d(a^2)/dt$ in the first part of the curve in figure 5, for small times, when a is approximately constant, is compared with the calculated value. The experimental rates are greater than those calculated from equation (23) by a factor of approximately 2. It is believed that this is due to the conditions of the experiment, for the larger drops are situated more at the end of the fibre, owing to their greater weight. A drop smaller in weight by a factor of 10 is more pear-shaped owing to the finite thickness of the fibre, and the neck represents a considerable proportion of the total surface. Since the sphere has the smallest surface for a given volume, it follows that the transference of the data from drops of mass 6×10^{-4} to 6×10^{-5} g. will be subject to an error of a geometrical nature, and the smaller drop will evaporate faster than the theory allows. It is probable that the geometrical factor remains approximately constant during a run, once the quartz fibre has been wet to a certain length, so that the function $\phi(a)$ is indeed linear with time. The time required to evaporate completely an isolated drop free from all attachment is best calculated from experiments on hanging drops over a small range of change in radius.

Runs on butyl stearate

In table 4 are summarized the results with butyl stearate, using drops which remained approximately constant in radius during the run. In figure 8, $-10^6\,ds/dt$ is plotted against $1/P$, and in figure 9, $q_0/q-\dfrac{1}{1+kt/(AaP)}$ is plotted against $1/P$, taking the collision radius, i.e. the sum of the effective radii of butyl stearate and

TABLE 4. BUTYL STEARATE AT 20°

mean a during run mm.	P mm. of Hg	$-10^6\times ds/dt$ cm.^{2} per min. observed	q_0/q	$-10^6\times ds/dt$ calc. from equation (25) cm.^{2} per min.
0·542	11·19	0·084	1	0·0848
0·544	4·77	0·190	1·034	0·192
0·543	2·33	0·345	1·167	0·366
0·545	0·743	0·977	1·292	0·959
0·533	0·309	1·489	2·038	1·467
0·536	0·238	1·732	2·440	1·647
0·553	0·164	1·833	3·121	1·875

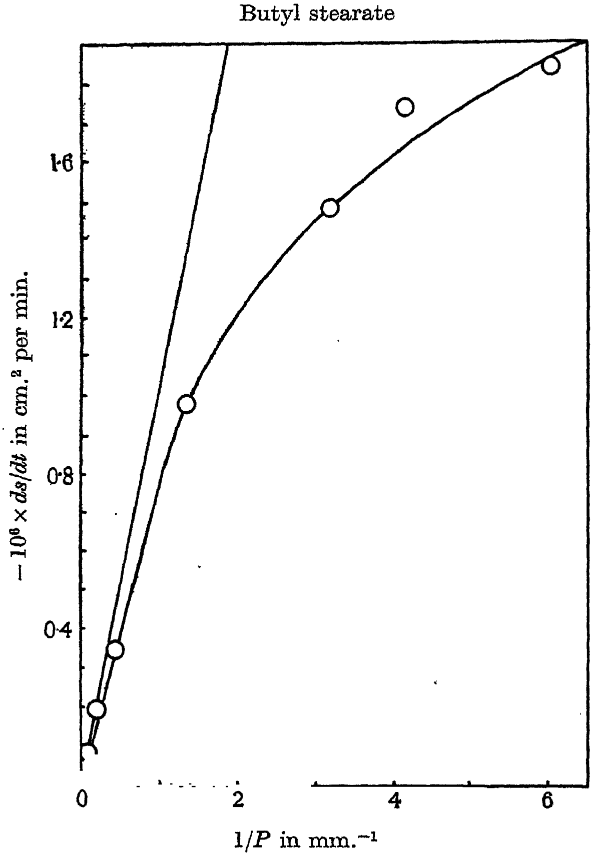


Figure 8

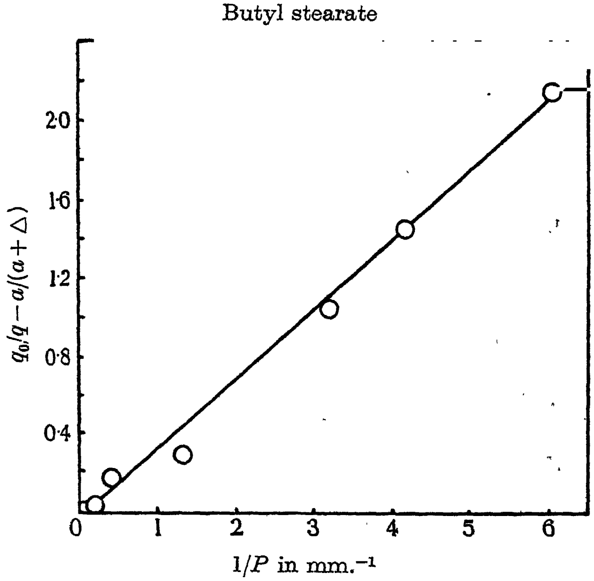


Figure 9

air molecules, as 17 A; linear plots of s versus time were obtained for all runs. The line in figure 8 has been calculated from the equation

$$-10^6 \times ds/dt = \frac{0{\cdot}980}{P}\left[\frac{1}{\dfrac{1{\cdot}957\times 10^{-2}}{aP}+\dfrac{1}{1+3{\cdot}337\times 10^{-4}/(aP)}}\right], \qquad (25)$$

where P is measured in mm. of mercury, a in cm., t in min., and s in cm.^{2}; the mean value of a was 0·542 mm.

IV. EXPERIMENTAL

The determination of the vapour pressure of dibutyl phalate and butyl stearate

The method was a differential modification of that due to Knudsen (1909). The ends of two glass tubes A and B were closed by sheet platinum caps fused to the glass (figure 10) and pierced by two circular holes 1·88 and 1·025 mm. in diameter. By means of side tubes the glass tubes could be half-filled with liquid and sealed off. A and B were made as nearly alike as possible.

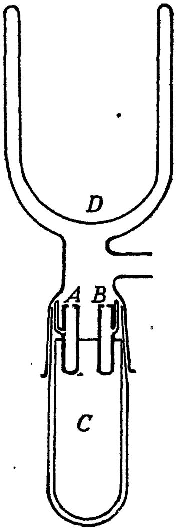


FIGURE 10

The tubes A and B were fitted into holes in the solid brass rod C, which was mounted inside a ground joint of the glass apparatus. D was a silvered Dewar flask which was kept full of liquid air. The lower portion of the apparatus was immersed in mercury at constant temperature.

The apparatus was evacuated by means of a mercury vapour diffusion pump, with a liquid air trap between pump and apparatus, for about an hour. Dry air was admitted and A and B were hung from opposite ends of a micro-analytical balance weighing to 10^{-6} g. The difference in weight between A and B was thus determined. A and B were replaced in the apparatus which was maintained at a high vacuum for several hours. The difference in weight between A and B was again determined. For A

$$p_\infty = \frac{W_1}{t}\left(\frac{RT}{M}\right)^{\frac{1}{2}}\frac{1}{A_1}\cdot\left[\sqrt{2\pi}+\frac{L_1}{R_1}\frac{3\sqrt{\pi}}{4\sqrt{2}}\right] \text{ (approx.)},$$

where p_∞ is the vapour pressure, W_1 the weight lost in time t, R the gas constant, M the molecular weight, A_1 the area of the hole and R_1 its radius, L_1 the length of the tube through which the vapour passes (i.e. thickness of platinum) (cf. Clausing 1932; Dushmann 1926). The calculation for the small correction due to the tube length is approximate. Hence $p_\infty = \frac{k_1 W_1}{tA_1}\left(1+\frac{x}{R_1}\right)$, where $k_1 = \left(\frac{2\pi RT}{M}\right)^{\frac{1}{2}}$ and x is a constant to allow for the tube resistance.

Similarly $$p_\infty = k_1\frac{W_2}{tA_2}\left(1+\frac{x}{R_2}\right).$$

Hence $$W_1 - W_2 = \frac{A_1 p_\infty t}{k_1}\left(1-\frac{x}{R_1}\right) - \frac{A_2 p_\infty t}{k_1}\left(1-\frac{x}{R_2}\right) \text{ approx.} \tag{26}$$

When $x = 0$ we get

$$W_1 - W_2 = p_\infty\frac{(A_1-A_2)t}{k_1} \quad \text{or} \quad p_\infty = \frac{k_1(W_1-W_2)}{(A_1-A_2)t}. \tag{27}$$

The second approximation gives

$$p_\infty = \frac{k_1(W_1-W_2)}{(A_1-A_2)t - \left(\frac{A_1}{R_1}-\frac{A_2}{R_2}\right)xt}. \tag{28}$$

$(W_1 - W_2)$ was equal to the difference in readings of the analytical balance, and A_1 and A_2 could be measured. Using very pure mercury as a calibrating liquid x could be determined and the x term was shown to be small, i.e. the conditions above were fulfilled.

Equation (28) may be rewritten

$$p_\infty = \frac{k_2}{(M)^{\frac{1}{2}}}\frac{(W_1-W_2)}{t}, \tag{29}$$

where k_2 is a constant determined using mercury and which can be used to find the vapour pressure of the organic liquids.

The mean results were, for 20°,

dibutyl phthalate: $p_\infty = 3\cdot05\times10^{-5}$ mm. of mercury,
butyl stearate: $p_\infty = 1\cdot27\times10^{-6}$ mm. of mercury.

V. The calculation of evaporation and diffusion coefficients

Using Langmuir's equation $-\frac{ds}{dt} = \frac{8\pi D m_2 c_0}{\rho}$ it is clear that the diffusion coefficient may be calculated from the runs on droplets in the high-pressure range if c_0 be replaced by p_0/kT, its value assuming the vapour to obey the perfect gas laws.

Hence
$$D \text{ at pressure } P = \left(-\frac{ds}{dt}\right)_{\text{press. } P} \frac{\rho k T}{8\pi p_0 m_2}. \tag{30}$$

The diffusion coefficient at atmospheric pressure may be calculated by simple proportion, since D is inversely proportional to P. The results of § III give, for 20° (putting $p_\infty = p_0$),

dibutyl phthalate:	diffusion coefficient at 760 mm. = 0·031	($\rho = 1{\cdot}048$),
butyl stearate:	diffusion coefficient at 760 mm. = 0·030	($\rho = 0{\cdot}867$).

This method of calculating diffusion coefficients from the Langmuir region of evaporation has been previously used in this Department by Mrs Childe (*née* Stevenson) for more volatile materials than those used in this work; a quartz spiral spring was used (unpublished Ph.D. thesis). Vapour pressures were determined by Knudsen's method. The differential method for vapour pressures, using two evaporating vessels, is new to this work.

The evaporation coefficient may be calculated using equation (12). The first term of the denominator within the brackets is $Y/Pav\alpha$, and its numerical value is known from equations (23) and (25), e.g. equation (23) gives $\frac{Y}{Pav\alpha} = \frac{2{\cdot}15 \times 10^{-2}}{P\alpha}$. Hence for dibutyl phthalate $\alpha = Y/(2{\cdot}15 \times 10^{-2} v)$. Since $Y = DP$ and D has already been determined for any value of P, Y may be calculated; for dibutyl phthalate Y is clearly $0{\cdot}031 \times 760$ in the units of pressure used, mm. of Hg. Hence α can be calculated. The results for the evaporation coefficient of dibutyl phthalate and butyl stearate at 20° are respectively 0·28 and 0·35.

VI. Conclusion and discussion

The work reported above on the evaporation of droplets at various air pressures leaves no doubt that Fuchs's equation represents the facts, and it is to be expected that the application of this equation to conditions other than those investigated, e.g. to conditions of very small radius and atmospheric pressure, will give an equally true picture of the evaporation behaviour of droplets.

Some biological and industrial applications are of interest. Thus evaporation from the pores of the skin of animals, or the stomata of plants will be enhanced by the smallness of these openings, for a given total area of evaporating water surface, provided that no interference due to vapour occurs between neighbouring openings. For evaporation from hemispherical caps of radius a it is seen that

$$\text{rate of evaporation per unit area} = \frac{Dc_0 m_2}{\dfrac{D}{\nu\alpha} + \dfrac{a^2}{a+\Delta}}.$$

$$\text{When } a \text{ is large, rate of evaporation per unit area} = \frac{Dc_0 m_2}{a_1}.$$

$$\text{Hence}\quad \left(\frac{\text{rate of evaporation per unit area for small drops}}{\text{rate of evaporation per unit area for large drops}}\right) = \frac{a_1}{\dfrac{D}{\nu\alpha} + \dfrac{a^2}{a+\Delta}}, \tag{31}$$

where $a_1 \gg a$. For water at atmospheric pressure of air $D/(\nu\alpha) = 0{\cdot}001$ approx., taking $\alpha = 0{\cdot}01$, $D = 0{\cdot}2$ approx. Hence the ratio in equation (31) becomes $\dfrac{a_1}{0{\cdot}001 + a}$ approx., which with $a_1 = 1$ cm. and $a = 10^{-4}$ cm. gives an increase by a factor of approx. 900; with $a_1 = 1$ cm. and $a = 10^{-3}$ cm. the factor is approx. 500. It should be noted that in this calculation no modification of the theory is introduced to allow for the presence of a liquid edge. Moreover, in practice these factors will be cut down by vapour interference between evaporating surfaces. Other applications are in the drying of foodstuffs by spraying, where there will be a large factor increasing the rate of evaporation even when the extra surface is allowed for, and the possibility of separating constituents of oils by evaporation from sprays.

We wish to express our thanks to the Ministry of Supply for permission to publish this work.

REFERENCES

Clausing 1932 *Ann. Phys., Lpz.*, **12**, 961.
Dushmann 1926 *Int. Crit. Tables*, **1**, 91.
Fuchs, N. 1934 *Phys. Z. Sowjet.* **6**, 225.
Knudsen 1909 *Ann. Phys., Lpz.*, **28**, 999.
Langmuir 1918 *Phys. Rev.* **12**, 368.
Whytlaw-Gray & Patterson 1932 *Smoke*. London: Arnold & Co.

Fluid motion between parallel planes. Dynamical stability

By D. Meksyn, D.Sc.

(*Communicated by G. Temple, F.R.S.—Received* 3 *April* 1945)

If U is the velocity of the mean motion the following main results are obtained:

1. The region where $U=c$, c being the wave velocity, is the source where vibrations are generated; i.e. the slowly varying vibrations give rise to large rapidly varying vibrations in passing through the critical point.
2. Curved profiles admit a periodic motion at sufficiently high Reynolds numbers.
3. Parabolic flow is unstable at high Reynolds numbers; i.e. an infinitely small disturbance is sufficient to break up such flow.

The critical Reynolds number is equal to $R=U_0h/\nu=6700$, and the corresponding wavelength is about three times the width of the channel (U_0 is the mean velocity at the axis, and h is the half-width of the channel).

1. INTRODUCTION AND SUMMARY

Many attempts have been made to prove that laminar flow is unstable at sufficiently high Reynolds numbers and that the subsequent motion becomes turbulent.

A critical survey of this question was given by Southwell & Chitty (1930). The problem of stability was also discussed by Noether (1926), Goldstein (1936) and Pekeris (1936, 1938).

Squire (1933) has shown that if the two-dimensional flow in a channel becomes unstable its critical Reynolds number is lower than the critical Reynolds number of a three-dimensional flow; and that, accordingly, it is sufficient to treat the problem of stability in two dimensions.

Heisenberg (1924), Tietjens (1925) and Tollmien (1929) considered several instances of laminar motion and found instability. The results, however, of Heisenberg and Tollmien are invalid, since their transformations of the non-viscous integrals through the critical point are incorrect. Tietjens discussed the special case of profiles consisting of broken straight lines; this eliminates the transformation of the non-viscous integrals. The transformation of viscous integrals was found numerically.

The problem of turbulent motion consists, in the first instance, in finding a solution of a fourth order differential equation containing a large parameter. The equation has two slowly varying, non-viscous integrals and two rapidly varying, viscous integrals. Because of the large parameter the integrals have to be expressed in an asymptotic form.

In order that the solution may satisfy the required boundary conditions the wave velocity has to be equal to the mean velocity at some region between the axis and the planes (Noether 1926, p. 234), with the result that the asymptotic expansions have a critical point.

The main mathematical difficulty of this problem is in finding the transformations of the integrals through the critical point.

In the present paper the flow between parallel planes is considered, and the equations of motion are solved under the condition that ϵ (3·15) and α^2 (2·6) are small, so that their squares and higher powers can be neglected.

If U is the mean motion the following main results are obtained:

(1) The region where $U = c$, c being the wave velocity, is the source where vibrations are generated; i.e. the slowly varying vibrations give rise to large, rapidly varying vibrations in passing through the critical region (§ 3).

(2) Curved profiles admit a periodic motion at sufficiently high Reynolds numbers (§ 5).

(3) Parabolic flow is unstable at high Reynolds numbers, i.e. an infinitely small disturbance is sufficient to break up such flow.

The critical Reynolds number is equal to $R = U_0 h/\nu = 6700$, and the corresponding wave-length is about three times the width of the channel (U_0 is the mean velocity at the axis, and h is the half-width of the channel) (§ 6).

2. Equations of motion

Consider the flow between two parallel planes, the direction of the x and y axes being parallel and normal to the planes respectively; the x direction is also parallel to the mean flow. The origin of co-ordinates is midway between the planes, the distance between the planes being equal to two.

The equations of motion are

$$\left.\begin{aligned} \frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\nabla^2 u,\\ \frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\nabla^2 v,\\ \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y} &= 0. \end{aligned}\right\} \tag{2·1}$$

Denoting as usual

$$u = \frac{\partial \psi}{\partial y}, \quad v = -\frac{\partial \psi}{\partial x}, \quad \zeta = \frac{\partial u}{\partial y}-\frac{\partial v}{\partial x} = \nabla^2\psi, \tag{2·2}$$

and eliminating p between equations (2·1), then

$$\frac{\partial \zeta}{\partial t}+u\frac{\partial \zeta}{\partial x}+v\frac{\partial \zeta}{\partial y} = \nu\nabla^2\zeta. \tag{2·3}$$

Put
$$u = U+u', \quad v = v', \quad \zeta = \frac{dU}{dy}+\zeta', \quad w = U/U_0, \tag{2·4}$$

where U is the mean motion, and it is a function of y only, and u', v' and ζ' depend also on time.

Substituting (2·4) in (2·3) and dropping the primes, then

$$\frac{\partial^2\zeta}{\partial x^2}+\frac{\partial^2\zeta}{\partial y^2}-\frac{1}{\nu}\left(\frac{\partial\zeta}{\partial t}+U\frac{\partial\zeta}{\partial x}+v\frac{\partial^2 U}{\partial y^2}\right)=\frac{1}{\nu}\left(u\frac{\partial\zeta}{\partial x}+v\frac{\partial\zeta}{\partial y}\right). \tag{2·5}$$

In the first approximation neglect the quadratic terms and put

$$\psi \sim e^{i(\beta t-\alpha x)}, \tag{2·6}$$

whence (2·5) becomes

$$\frac{d^4\psi}{dy^4}+\left[-2\alpha^2-\frac{i(\beta-\alpha U)}{\nu}\right]\frac{d^2\psi}{dy^2}+\left[\alpha^4+\frac{i\alpha^2(\beta-\alpha U)}{\nu}-\frac{i\alpha}{\nu}\frac{d^2U}{dy^2}\right]\psi=0. \tag{2·7}$$

The solution ψ has to satisfy the boundary conditions

$$\psi=0=\psi', \quad \text{at} \quad y=\pm 1. \tag{2·8}$$

The equation (2·7) has four integrals. Two of them are rapidly varying functions; they are due to viscosity and are only sensible near the walls. The other pair depends very little on viscosity for the greater part of the channel; they change slowly across the channel, and are mainly responsible for the turbulent velocities as usually observed.

The equation (2·7) has to be solved for real values of α and β; it is assumed that $\beta-\alpha U$ vanishes at some point between the axis and the boundary, and that $w''\neq 0$ at that point.

Since $w-c$ is an even function in y, the equation (2·7) has two even and two odd solutions, and only the range of y between zero and unity need to be considered.

In the following only the even solutions will be discussed.

3. NON-VISCOUS INTEGRALS

It is convenient to transform the equation (2·7) as follows. Let the half-width of the channel be h; put

$$x'=\frac{x}{h},\quad y'=\frac{y}{h},\quad u'=\frac{u}{U_0},\quad v'=\frac{v}{U_0},\quad t'=\frac{tU_0}{h}, \tag{3·1}$$

where U_0 is the mean velocity along the axis of the channel. The equation (2·7) is transformed, after dropping the primes, into

$$(w-c)(\psi''-\alpha^2\psi)-w''\psi=\frac{i}{\alpha R}(\psi^{\text{iv}}-2\alpha^2\psi''+\alpha^4\psi),$$

where
$$R=\frac{U_0 h}{\nu},\quad c=\beta/\alpha U_0,\quad w=\frac{U}{U_0},\quad w-c=\frac{1}{U_0}(U-\beta/\alpha). \tag{3·2}$$

Consider first the equation

$$(w-c)\psi''-w''\psi=0. \tag{3·3}$$

One integral, say, ψ_2 is found by inspection

$$\psi_2=w-c, \tag{3·4}$$

and, applying the method of variation of parameters, the second integral is found as

$$\psi_1 = (w-c)\int \frac{dy}{(w-c)^2}. \tag{3·5}$$

Assuming that α^2 and $1/\alpha R$ are sufficiently small the solution of (3·2) can be expanded in powers of α^2 and $1/\alpha R$ by the method of successive approximations.

Rewriting the equation (3·2) as

$$\psi'' - \frac{w''}{w-c}\psi = F, \tag{3·6}$$

a particular solution of (3·6) is

$$\psi = (w-c)\int^y \frac{\int^y (w-c)\,F\,dy}{(w-c)^2}\,dy,$$

and a simple computation leads to

$$\left.\begin{aligned} \psi_1 &= (w-c)\int_0^y \frac{dy}{(w-c)^2}\left\{1+\alpha^2\int_0^y dy(w-c)^2\int_0^y \frac{dy}{(w-c)^2}+\dots\right. \\ &\qquad\left. +\frac{i}{\alpha R}\left[\frac{d^3}{dy^3}(w-c)\int_0^y \frac{dy}{(w-c)^2}+\dots\right]\right\}, \\ \psi_2 &= (w-c)\left\{1+\alpha^2\int_0^y \frac{dy}{(w-c)^2}\int_0^y (w-c)^2\,dy+\dots+\frac{i}{\alpha R}\int_0^y \frac{w'''+\dots}{(w-c)^2}\,dy\right\}, \end{aligned}\right\} \tag{3·7}$$

where the integrals denote successive integrations and not products (Heisenberg 1924, p. 586), and only terms up to the order α^2 and $1/\alpha R$ are retained.

The expressions (3·7) do not depend, in the first approximation, on Reynolds's number, i.e. on viscosity; they change slowly across the channel between the critical points; they can be, accordingly, termed as non-viscous, slowly varying, integrals.

Since $w-c$ is an even function in y, ψ_1 and ψ_2 are odd and even in y respectively.

It can be shown (Noether 1926, p. 234) that, unless $w-c$ vanishes somewhere between the axis and the boundary, the equations (2·7) and (2·8) cannot be satisfied for real values of c and α; it is, therefore, assumed that

$$w-c=0 \quad \text{at} \quad y=y_0. \tag{3·8}$$

There will be also a root at $y=-y_0$.

It will be shown below that some of the terms in ψ_1 and ψ_2 become infinite at y_0; these terms are not valid near y_0, and they change their form in passing through this point; on the other hand, the terms which remain finite, i.e. the regular terms, preserve their form throughout the region.

The singular components of the integrals ψ_1 and ψ_2 can therefore be treated as asymptotic approximations valid only between the axis and up to a certain distance from y_0; the latter point is a critical point of these asymptotic expansions, although it is an ordinary point of the equation (2·7).

Transformation through the critical point

To fix ideas consider the integral ψ_2 even in y. The first term, i.e. $w-c$, is regular, and it preserves its form in passing through the critical point y_0.

Consider now the term proportional to α^2, namely,

$$(w-c)\int_0^y \frac{dy}{(w-c)^2}\int_0^y (w-c)^2\,dy. \tag{3·9}$$

Transforming (3·9) to a new variable z, where

$$z = y-y_0, \tag{3·10}$$

one obtains, by splitting up (3·9) in to two parts,

$$(w-c)\int_{-y_0}^{z} \frac{dz}{(w-c)^2}\int_{-y_0}^{0} (w-c)^2\,dz + (w-c)\int_{-y_0}^{z} \frac{dz}{(w-c)^2}\int_0^z (w-c)^2\,dz. \tag{3·11}$$

The second term is regular (it remains, together with its derivatives, finite within our range), and it need not be considered.

Expanding $(w-c)^2$ round the point $z=0$ in the denominator of the first term (3·11) and retaining only two terms, then

$$p(w-c)\int_{-y_0}^{z} \frac{1}{w_0'^2 z^2}\left(1-\frac{w_0''}{w_0'}z\right)dz, \quad p = \int_{-y_0}^{0} (w-c)^2\,dz = \int_0^{y_0} (w-c)^2\,dy, \tag{3·12}$$

where w_0' and w_0'' are derivatives with respect to y at $z=0$; whence the singular term is

$$-p\frac{w_0''}{w_0'^3}(w-c)\log(y_0-y), \tag{3·13}$$

and the regular component of ψ_2 up to the order α^2 inclusive is

$$\left.\begin{aligned} \psi_2 &= (w-c)+\alpha^2\left\{(w-c)\int_0^y \frac{dy}{(w-c)^2}\int_0^y (w-c)^2\,dy + p\frac{w_0''}{w_0'^3}(w-c)\log(y_0-y)\right\}, \\ p &= \int_0^{y_0} (w-c)^2\,dy. \end{aligned}\right\} \tag{3·14}$$

The above regular component of ψ_2 preserves its form, and it is only necessary to find how the term (3·13) is expressed round the critical point.

To that end the equation (3·2) is transformed to a new variable η, where

$$\epsilon\eta = z = y-y_0, \quad \epsilon = (w_0'\alpha R)^{-\frac{1}{3}}, \tag{3·15}$$

and the solution is expanded in powers of the small parameter ϵ.

The singular term (3·13), except for a constant, is

$$\left(w_0'z+\frac{w_0''z^2}{1.2}+\dots\right)\log(-z), \tag{3·16}$$

or, expressed in the new variable η,

$$\left(w_0'\epsilon\eta+\frac{w_0''\epsilon^2\eta^2}{1.2}+\dots\right)\log(-\epsilon\eta). \tag{3·17}$$

Since ϵ is usually very small and η, as it will be shown below, is moderate within the asymptotic range and does not depend on the value of ϵ, the first term only is retained in (3·17); i.e. one has to find the transformation of

$$\eta\log\eta. \tag{3·18}$$

Now, ϵ (3·15) is negative since w_0' is negative; accordingly the following relations hold good (3·15):

$$z=y-y_0\lesseqgtr 0;\quad \eta\gtrless 0, \tag{3·19}$$

and the problem is reduced to the following to find an integral of (3·2) which for positive η asymptotically tends to (3·18) to the order of ϵ.

Although z is small, ϵ being very small, η can be made large enough to lie within the asymptotic range of the integral obtained. In fact, it will be shown that $\eta = 4$ satisfies the required conditions.

The solution of this rather involved problem is given elsewhere (Meksyn); the final result is as follows:

The required integral is $T(\eta)+\eta$, where

$$T(\eta)=-0{\cdot}813+0{\cdot}469i+\left(-0{\cdot}751-\frac{\pi i}{6}\right)\eta-\sum_{n=1}^{\infty}\frac{1}{3}\Gamma\left(\frac{n}{3}\right)\frac{(-3^{\frac{1}{3}}e^{\pi i/6})^n\,\eta^{n+1}}{(n+1)!} \tag{3·20}$$

and $T(\eta)+\eta\rightarrow\eta\log\eta$, when $\eta>0$.

When η is negative and large, $T(\eta)$ tends to

$$T(\eta)\rightarrow\frac{\pi^{\frac{1}{2}}e^{-3\pi i/8}\exp\left[\frac{2}{3}e^{\pi i/4}(-\eta)^{\frac{3}{2}}\right]}{(-\eta)^{\frac{5}{4}}}. \tag{3·21}$$

The above expression is, however, identical with that of the viscous integral (4·10).

Therefore the following very important result accrues: that *a component of the non-viscous integral* which is almost equal to zero at the axis *becomes transformed into a viscous integral in passing through the critical region.*

The physical interpretation of this result is that *the critical region is a source where vibrations are generated*; that explains how turbulent motion is possible.

Viscous forces generate vibrations within the critical region and these vibrations are then destroyed near the axis of the channel; when equilibrium is established between these two forces turbulent motion can persist.*

Below are given the values of $T(\eta)+\eta$ and its derivatives for a set of values of η, together with the corresponding curves.

---

* It was pointed out by G. I. Taylor (*Phil. Trans.* (1915) A, **215**, 23–26) that vorticity may make inviscid flow unstable.

Taylor (*Proc. Roy. Soc.* (1935), A, **151**, 455–464) has also proved that near the walls of a channel more turbulent energy is generated than destroyed, and that the opposite takes place near the axis.

TABLE 1

η	$T(\eta)+\eta$	$\eta\log\eta$	$T'(\eta)+1$	$\log\eta+1$
4·0	5·544 + 0·020i	5·545	2·388 − 0·009i	2·386
3·5	4·378 + 0·026i	4·385	2·255 − 0·014i	2·253
3·0	3·292 + 0·035i	3·296	2·103 − 0·021i	2·099
2·5	2·283 + 0·048i	2·291	1·927 − 0·033i	1·916
2·0	1·371 + 0·070i	1·386	1·717 − 0·054i	1·693
1·5	0·592 + 0·137i	0·608	1·461 − 0·089i	1·405
1·0	−0·039 + 0·166i	0	1·146 − 0·159i	1·000
0·5	−0·558 + 0·273i	−0·347	0·750 − 0·285i	0·307
0	−0·813 + 0·469i	0	0·249 − 0·524i	$-\infty$
−0·5	−0·789 + 0·831i	—	−0·365 − 0·970i	—
−1·0	−0·433 + 1·502i	—	−1·056 − 1·794i	—
−1·5	0·236 + 2·699i	—	−1·649 − 3·243i	—
−2·0	1·120 + 4·901i	—	−1·658 − 5·612i	—
−2·5	1·640 + 8·519i	—	−0·015 − 8·954i	—
−3·0	0·552 + 13·871i	—	5·141 − 12·304i	—
−3·5	−4·442 + 20·23i	—	15·862 − 11·974i	—
−4·0	−16·12 + 23·81i	—	31·06 + 0·758i	—

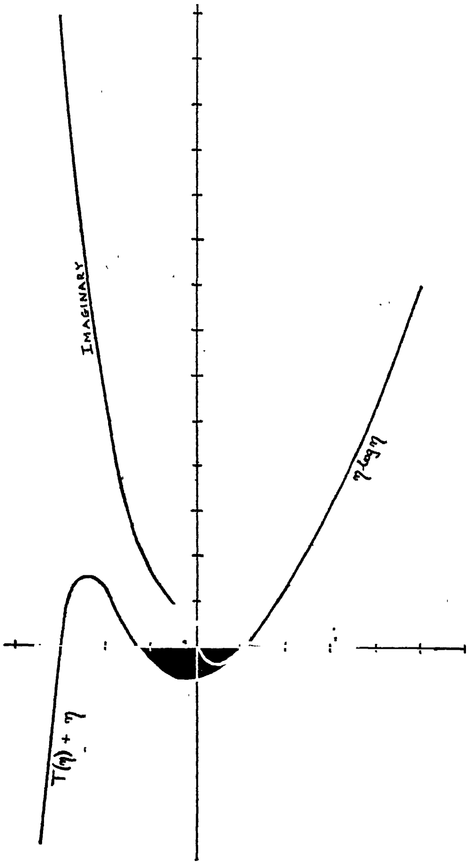


FIGURE 1. $T(\eta)+\eta$ and $\eta\log\eta$.

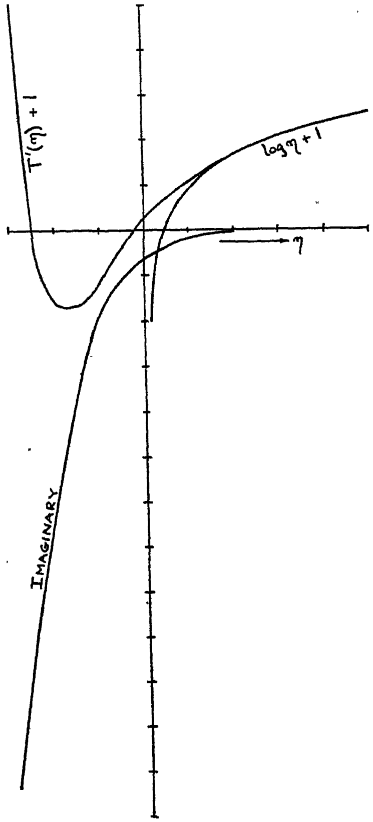


FIGURE 2. $T'(\eta)+1$ and $\log\eta+1$.

It can be seen that the asymptotic expression $\eta \log \eta$ is valid even very close to the critical point.

The solution was pursued up to the first power of ϵ; there is no difficulty in finding the corresponding transformation for terms of higher power in ϵ; they are not needed, however, in the present case.

Since $$(y-y_0)\log(y_0-y) = z\log(-z) = \epsilon\eta(\log\eta + \log|\epsilon|), \tag{3·22}$$

the corresponding transformation is (3·20)

$$z\log(-z) \rightarrow z + z\log|\epsilon| + \epsilon T(\eta) \quad (z<0,\ \eta>0), \tag{3·23}$$

and the even solution becomes

$$\left.\begin{aligned}
\psi_2 = (w-c) + \alpha^2\Big\{&(w-c)\int_0^y \frac{dy}{(w-c)^2}\int_0^y (w-c)^2\,dy + p\frac{w_0''}{w_0'^3}(w-c)\log(y-y_0)\Big\} \\
&-\alpha^2 p\frac{w_0''}{w_0'^2}\epsilon\Big\{-0{\cdot}813 + 0{\cdot}469i + \left(0{\cdot}249 + \log|\epsilon| - \frac{\pi i}{6}\right)\eta \\
&-\sum_{n=1}^{\infty}\tfrac{1}{3}\Gamma\left(\frac{n}{3}\right)\frac{(-3^{\frac{1}{3}}e^{\pi i/6})^n\,\eta^{n+1}}{(n+1)!}\Big\}, \\
p = \int_0^{y_0}&(w-c)^2\,dy.
\end{aligned}\right\} \tag{3·24}$$

The expression in the last bracket on the right-hand side (3·24) is equal to (3·22)

$$\eta + \eta\log|\epsilon| + T(\eta).$$

4. Rapidly varying (viscous) integrals

To find the rapidly varying integrals of (2·7) assume

$$\psi = Ae^{i\xi}, \tag{4·1}$$

where A is a slowly varying function of y, and ξ includes the large parameter (αR), which is contained in $(\beta-\alpha U)/\nu$; α is assumed small.

Differentiating (4·1), substituting in (2·7) and equating to zero the two terms of the highest order, then, by rather tedious calculation,

$$\xi'^2 = -\frac{i(\beta-\alpha U)}{\nu}, \quad \xi = \int_{y_0}^{y}\xi'\,dy, \quad A = \frac{\text{const.}}{\xi'^{\frac{5}{2}}}, \tag{4·2}$$

and the two rapidly varying integrals are

$$\phi_3 = \frac{1}{\xi'^{\frac{5}{2}}}\sin\xi, \quad \phi_4 = \frac{1}{\xi'^{\frac{5}{2}}}\cos\xi, \tag{4·3}$$

where the letters ϕ_3 and ϕ_4 are used instead of ψ_3 and ψ_4, the latter being reserved for odd and even functions in y.

It can be seen from (4·2) that the integrals depend on R, i.e. viscosity, and for large R they are rapidly varying functions of y.

Now to construct from (4·3) an even solution in y. The coefficients in (2·7) are even functions in y, and the equation has accordingly two odd and two even integrals.

Let ψ_4 be an even solution different from ψ_2; since $y = 0$ is an ordinary point of the equation, ψ_4 will be completely determined if its value, together with its derivatives up to the third order inclusive, are specified at $y = 0$.

Since ψ_4 is an even function it satisfies the conditions

$$\psi_4 = \text{finite}, \quad \psi_4' = 0 = \psi_4''', \quad \psi_4'' = \text{finite} \quad \text{at} \quad y = 0. \tag{4·4}$$

Assume, accordingly (3·7), (4·3),

$$\psi_4 = a\psi_1 + b\psi_2 + c\phi_3 + d\phi_4, \tag{4·5}$$

where a, b, c and d are constants.

Now, ψ_2 is an even function, and it can be neglected if one is not interested in the precise value of $\psi''(0)$; there remains now only three equations. It can be shown *a posteriori* that if the term $a\psi_1$ is disregarded the function ψ_4, thus derived, will satisfy all the three remaining conditions (4·4).

Assume accordingly at $y = 0$ that

$$\psi_4 = c\phi_3 + d\phi_4 = \text{finite}, \quad \psi_4' = c\phi_3' + d\phi_4' = 0. \tag{4·6}$$

whence substituting the values of ϕ_3 and ϕ_4 (4·3) and solving with respect to c and d one finds as the dominant term

$$\psi_4 = \frac{1}{\xi'^{\frac{5}{2}}} \cos(\xi - \xi_0), \quad \xi_0 = \int_{y_0}^{0} \xi' \, dy. \tag{4·7}$$

The expression (4·7) is valid between $y = 0$ and the critical point $y = y_0$, but not too close to it, since at that point ψ_4 becomes infinite ($\xi' = 0$); it changes its form in passing through the critical point.

The derivation of the continuation of (4·7) through the critical point is very involved, and it will be given elsewhere (Meksyn); here only the final result is quoted.

The function ψ_4 (4·7) becomes near $\eta = 0$ (3·15)

$$\left.\begin{aligned} \frac{\psi_4}{a} &= 0·548 + 0·548i - (0·965 + 0·258i)\,\eta \\ &\quad + \int_0^\eta \int_0^\eta \eta^{\frac{1}{2}} J_{\frac{1}{3}}\left(\tfrac{2}{3} e^{9\pi i/4} \eta^{\frac{3}{2}}\right) d\eta^2 + (-0·5 + 0·866i) \int_0^\eta \int_0^\eta \eta^{\frac{1}{2}} J_{-\frac{1}{3}}\left(\tfrac{2}{3} e^{9\pi i/4} \eta^{\frac{3}{2}}\right) d\eta^2, \\ a &= \tfrac{1}{3} \pi^{\frac{1}{2}} e^{5\pi i/12} e^{-i\xi_0} \,|\epsilon|^{\frac{5}{2}}, \end{aligned}\right\} \tag{4·8}$$

where η is given by (3·15) and J is the well-known Bessel function, the series expansion being

$$J_\mu(x) = \sum_{m=0}^{\infty} \frac{(-)^m \left(\frac{1}{2}x\right)^{\mu+2m}}{m!\,\Gamma(\mu+m+1)}. \tag{4·9}$$

For large negative values of η, i.e. between the critical point and the boundary, ψ_4 tends asymptotically to

$$\frac{\psi_4}{a} = \frac{3}{2\sqrt{\pi}} e^{-\pi i/24} \eta^{-\frac{7}{4}} \exp\left[\tfrac{2}{3} e^{\pi i/4} (-\eta)^{\frac{3}{2}}\right], \tag{4·10}$$

which is of the same form as (3·21); this bears out the previous statement that the slowly varying solution tends to the rapidly varying one in passing through the critical point.

In table 2, ψ_4 and ψ_4' are given for several values of η; they are also plotted in figure 3.

TABLE 2

η	ψ_4	ψ_4'
4·0	—	$-0·002+\ 0·012i$
3·5	—	$0·056+\ 0·033i$
3·0	$-0·011-0·006i$	$0·030-\ 0·003i$
2·5	$-0·030+0·002i$	$0·045-\ 0·037i$
2·0	$-0·051+0·036i$	$0·034-\ 0·105i$
1·5	$-0·053+0·113i$	$-0·038-\ 0·201i$
1·0	$0·005+0·237i$	$-0·214-\ 0·293i$
0·5	$0·182+0·395i$	$-0·520-\ 0·330i$
0	$0·548+0·548i$	$-0·965-\ 0·258i$
−0·5	$1·170+0·632i$	$-1·549-\ 0·054i$
−1·0	$2·123+0·587i$	$-2·288+\ 0·240i$
−1·5	$3·493+0·400i$	$-3·236+\ 0·480i$
−2·0	$5·404+0·166i$	$-4·454+\ 0·360i$
−2·5	$7·980+0·188i$	$-5·852-\ 0·658i$
−3·0	$11·185+1·099i$	$-6·780-\ 3·329i$
−3·5	$14·372+3·883i$	$-5·309-\ 8·145i$
−4·0	$15·368+9·458i$	$2·430-13·511i$

As can be seen from these data, ψ_4 and ψ_4' tend rapidly to zero for η positive; in fact, a detailed examination of the solution (4·8) proves the above conclusion.

A viscous solution which vanishes for positive and large η was evaluated numerically by Tietjens (1925, p. 217), and a comparison can be made with the present analytical results.

Tietjens evaluated ψ_4/ψ_4' for a set of negative values of η; it is given by his constant $-D$; Tietjens's equation, however, is conjugate to that given here; accordingly the sign of the imaginary term in $-D$ has to be changed.

The corresponding values are given in table 3.

TABLE 3

$-\eta$	ψ_4/ψ_4'	Tietjens
0	$-0·670-0·387i$	$-0·702-0·425i$
0·5	$-0·769-0·381i$	$-0·785-0·411i$
1·0	$-0·892-0·350i$	$-0·920-0·389i$
1·5	$-1·038-0·278i$	$-1·043-0·297i$
2·0	$-1·202-0·135i$	$-1·206-0·147i$
2·5	$-1·350+0·119i$	$-1·357+0·108i$
3·0	$-1·393+0·522i$	$-1·400+0·515i$
3·5	$-1·142+1·020i$	$-1·180+1·130i$
4·0	$-0·480+1·224i$	$-0·460+1·250i$

The agreement is very satisfactory, since the numerical method, especially in the present case, can give only approximate results.

For very large negative η, the ratio ψ_4/ψ_4' tends to

$$-\frac{e^{-\pi i/4}}{|\eta|^{\frac{1}{2}}},$$

as can be seen from (4·10).

In figure 4 the real and imaginary parts of ψ_4/ψ_4' are plotted against each other and the variable η.

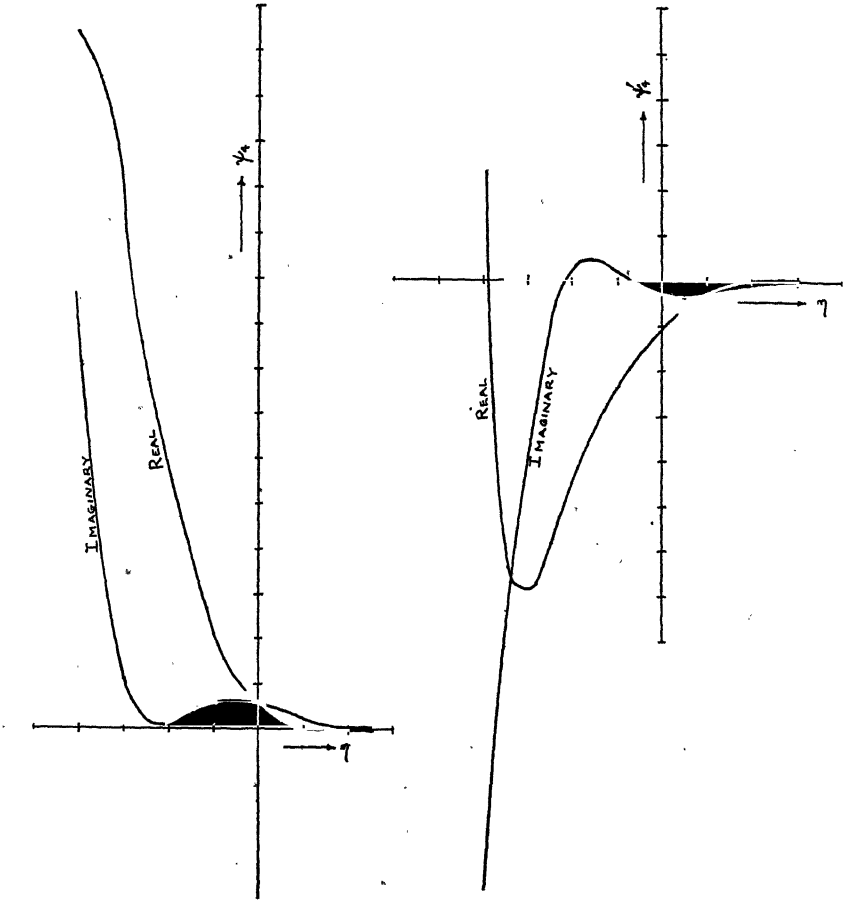


Figure 3

5. Fundamental equation. Critical R.

Only the even solutions of (2·7) will be considered; accordingly let

$$\psi = A\psi_2 + B\psi_4, \tag{5·1}$$

where A and B are constants.

The boundary conditions are

$$\psi = 0 = \psi' \quad \text{at} \quad y = 1. \tag{5·2}$$

Similar boundary conditions at $y = -1$ will also be satisfied since ψ is an even function in y.

From (5·1) and (5·2) follows the fundamental equation

$$\frac{\psi_2}{\psi_2'} = \frac{\psi_4}{\psi_4'} \quad \text{at} \quad y = 1. \tag{5·3}$$

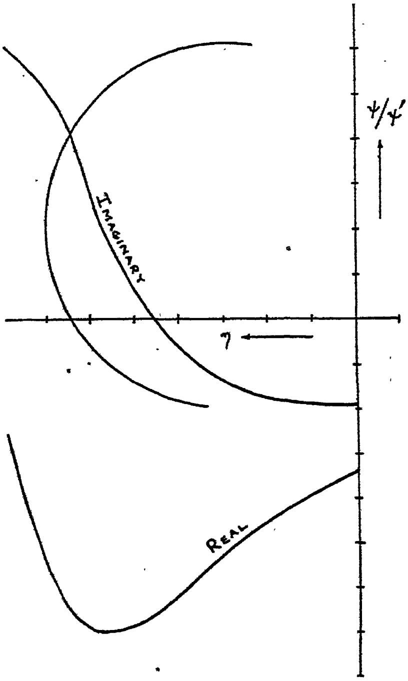


Figure 4

The functions ψ_2 and ψ_4 are given by (3·24), (4·8) and tabulated in tables 1–3.

It will be assumed below that the differentiation of ψ_2 and ψ_4 is taken with respect to η and not z, since that simplifies somewhat the computations.

Put for brevity

$$\psi_2 = A_1 + \alpha^2 A_2 + i\alpha^2 B_2, \quad \frac{\psi_4}{\psi_4'} = k + il, \tag{5·4}$$

where $A_1 = w - c$ and $\alpha^2 A_2$, $\alpha^2 B_2$ are the real and imaginary parts of the remaining terms in ψ_2; whence (5·3) becomes, after separating real and imaginary terms,

$$A_1 - A_1' k = \alpha^2\{-A_2 + A_2' k - B_2' l\}, \quad -A_1' l = \alpha^2\{-B_2 + A_2' l + B_2' k\}. \tag{5·5}$$

Eliminating α^2 one finds after few transformations

$$-A_1 B_2 + (A_1 A_2' - A_1' A_2)\,l + (A_1 B_2' + A_1' B_2)\,k - A_1' B_2'(k^2+l^2) = 0. \tag{5·6}$$

Eliminating A_1' between (5·5) then

$$\alpha^2 = \frac{A_1 l}{-A_2 l + B_2 k - B_2'(k^2+l^2)}. \tag{5·7}$$

A_1' was eliminated since for an arbitrary profile its value is rather more difficult to estimate than that of A_1.

The equations (5·6) and (5·7) have to be solved for real values of α and c which correspond to the wave-length and the frequency of the corresponding vibrations.

Since, however, $\qquad w - c = 0 \quad \text{at} \quad y = y_0,$

the quantity c is determined if the position of the critical point is known.

The equations (5·6) and (5·7) depend also on the quantity

$$\frac{w_0'}{w_0'' \epsilon}. \tag{5·8}$$

The equations are therefore solved in the following way. Assume that the boundary is at the distance $\eta = 0$, $\eta = -1$, ... from the critical point, and, making use of tables 1–3, (5·6) is evaluated for a set of negative values of η and the corresponding value of $w_0'/w_0''\epsilon$ is found; α^2 is then obtained from (5·7).

The computations are straightforward except for the following; A_2 includes a component proportional to (3·24)

$$A_0 = (w-c)\int_0^y \frac{dy}{(w-c)^2}\int_0^y (w-c)^2\,dy + p\,\frac{w_0''}{w_0'^3}(w-c)\log(y_0 - y), \tag{5·9}$$

and it will also be necessary to evaluate

$$A_1 A_0' - A_1' A_0 = \epsilon \int_0^1 (w-c)^2\,dy + p\,\frac{w_0''}{w_0'}\epsilon^2 \eta, \quad p = \int_0^{y_0} (w-c)^2\,dy. \tag{5·10}$$

To evaluate p (5·10) some assumption has to be made as to the expression for w, and the value of y_0; as a first approximation assume that

$$w - c = 1 - y^2, \quad y_0 = 1, \tag{5·11}$$

whence

$$p = \int_0^1 (1-y^2)^2\,dy = \frac{8}{15}. \tag{5·12}$$

The assumption $y_0 = 1$ is rather crude and it leads to somewhat high values for p; it does not affect greatly the value obtained for $w_0'/w_0''\epsilon$, because p drops out from the equation (5·6); it affects, however, the value of α.

In evaluating (5·9) only the dominant term is retained, whence

$$A_0 = -\frac{8}{15 w_0'}. \tag{5·13}$$

To elucidate the method of computation and the approximations involved one of the results in the table will be worked out briefly; namely, the case $\eta = -3$.

From (3·24) and table 1

$$\left.\begin{aligned} A_1 &= (w-c) = w_0'z = w_0'\epsilon\eta = -3w_0'\epsilon,\\ A_2 &= A_0 - \frac{w_0''\epsilon}{w_0'^2}p[0{\cdot}552 - 3\log|\epsilon|],\\ B_2 &= -13{\cdot}871\,\frac{w_0''\epsilon}{w_0'^2}p,\\ A_1' &= w_0'\epsilon,\\ A_2' &= A_0' - \frac{w_0''\epsilon}{w_0'^2}[5{\cdot}140 + \log|\epsilon|],\\ B_2' &= 12{\cdot}304\,\frac{w_0''\epsilon}{w_0'^2}p. \end{aligned}\right\} \tag{5·14}$$

Since

$$p = \int_0^{y_s}(w-c)^2\,dg \approx \int_0^1 (1-y)^2\,dy,$$

then (5·10)

$$\left.\begin{aligned} (w-c)A_0' - w_0'\epsilon A_0 &= p\left(\epsilon - 3\frac{w_0''}{w_0}\epsilon^2\right),\\ A_1B_2 &= 41{\cdot}612p\frac{w_0''\epsilon^2}{w_0'},\\ A_1A_2' - A_1'A_2 &= p\left(\epsilon + 12{\cdot}974\frac{w_0''\epsilon^2}{w_0'}\right),\\ A_1B_2' + A_1'B_2 &= -50{\cdot}783p\frac{w_0''\epsilon^2}{w_0'},\\ A_1'B_2' &= 12{\cdot}304p\frac{w_0''}{w_0'}\epsilon^2, \end{aligned}\right\} \tag{5·15}$$

whence substituting (5·15) in (5·6) together with the values of k and l from table 3 the results in table 4 are obtained for $\eta = -3$.

Table 4 and figure 5 give the results of these computations; for the sake of brevity it is denoted

$$\alpha^2 = \frac{Nw_0'^2\epsilon}{\Delta}. \tag{5·16}$$

TABLE 4

$-\eta$	$\psi_4/\psi_4' = k+il$	$w_0'/w_0''\epsilon$	N	Δ
0	$-0{\cdot}670-0{\cdot}387i$	0·81	—	—
0·5	$-0{\cdot}769-0{\cdot}381i$	0·90	0·190	$-0{\cdot}070+0{\cdot}112\log\|\epsilon\|$
1·0	$-0{\cdot}892-0{\cdot}350i$	1·90	0·350	$-0{\cdot}230+0{\cdot}098\log\|\epsilon\|$
1·5	$-1{\cdot}038-0{\cdot}278i$	3·95	0·417	$-0{\cdot}284+0{\cdot}056\log\|\epsilon\|$
2·0	$-1{\cdot}202-0{\cdot}135i$	14·36	—	—
2·5	$-1{\cdot}350+0{\cdot}119i$	−32·53	−0·297	$0{\cdot}141+0{\cdot}005\log\|\epsilon\|$
2·8	$-1{\cdot}397+0{\cdot}364i$	−19·22	—	—
3·0	$-1{\cdot}393+0{\cdot}522i$	−16·61	−1·566	$0{\cdot}523+0{\cdot}049\log\|\epsilon\|$
3·5	$-1{\cdot}142+1{\cdot}020i$	−20·45	−3·559	$0{\cdot}790+0{\cdot}093\log\|\epsilon\|$
4·0	$-0{\cdot}480+1{\cdot}224i$	−35·60	—	—

In figure 5 the quantity $w_0'/w_0''\epsilon$ is plotted against η; it starts from the small value of 0·81 and increases to ∞ when $l = 0$ which is at $\eta = -2{\cdot}3$; it then starts from $-\infty$, reaches its maximum value of about -16 at $\eta \eqsim -3$, and then tends to $-\infty$.

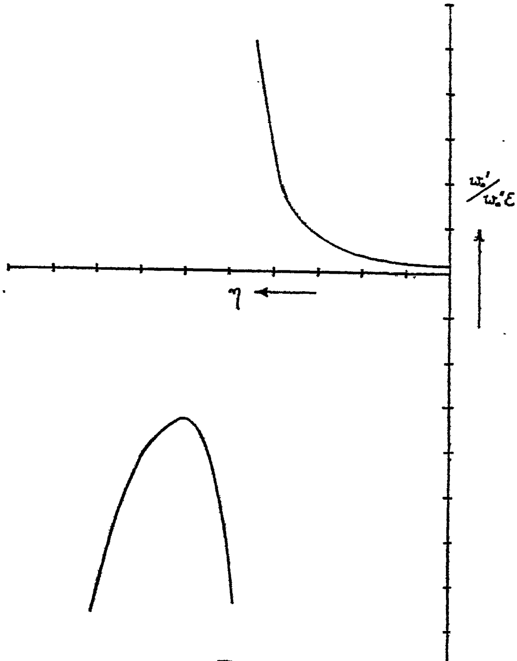


FIGURE 5

Very important conclusions can be derived from table 4 and figure 5.

The existence of a solution depends on the value of the quantity $w_0'/w_0''\epsilon$; the above result can be applied to any profile which does not change very rapidly and is decreasing from the axis to the boundary; for such profile w_0' is negative, and, whatever the value of w_0'/w_0'' may be, whether positive or negative, a quantity ϵ can be found sufficiently small, so that $w_0'/w_0''\epsilon$ lies on the curve, either on the positive or on the negative branch of it.

The following conclusion results:

Provided that Reynolds's number is sufficiently high (i.e. ϵ small) *curved profiles admit a periodic motion.*

It will be shown below that the flow near the boundary is largely parabolic, even in the case of turbulence; and it would be reasonable to assume that for the actual profiles w_0'/w_0'' is positive; accordingly, the lower curve is of greater importance, and from this it follows that a solution is only possible if

$$\left|\frac{w_0'}{w_0''\epsilon}\right| > 16,$$

i.e. if

$$R > \frac{16^3\,|w_0''|^3}{\alpha w_0'^4}, \qquad (5{\cdot}17)$$

which shows that generally speaking each profile admits a periodic solution only above a certain Reynolds number; the least Reynolds number corresponds to the vertex of the lower curve.

It should be emphasized that the above results are valid so long as α is not much larger than unity. There is, however, no difficulty in applying the same method to higher values of α; that might entail the consideration of terms of higher powers in ϵ, i.e. terms like $z^n \log(-z)$ when $n > 1$; the method of integration given here can be easily generalized for the latter case.

6. PARABOLIC FLOW

The critical Reynolds number when laminar motion becomes unstable corresponds to the vertex of the negative branch in figure 5.

(*a*) *First approximation.* As a first approximation assume, accordingly (figure 5),

$$\frac{w_0'}{w_0''\epsilon} = -16{\cdot}6, \quad \eta = -3. \tag{6·1}$$

The critical point is found from

$$1 - y_0 = \eta\epsilon = -3\epsilon. \tag{6·2}$$

For parabolic flow

$$w = 1 - y^2, \tag{6·3}$$

whence

$$\frac{w_0'}{w_0''\epsilon} = \frac{y_0}{\epsilon} = \frac{1+3\epsilon}{\epsilon} = -16{\cdot}6, \tag{6·4}$$

and solving (6·4)

$$\epsilon = -\frac{1}{19{\cdot}6}, \quad y_0 = 0{\cdot}85, \tag{6·5}$$

the corresponding value of α is (table 4)

$$\alpha = 0{\cdot}78,$$

whence

$$R = \frac{19{\cdot}6^3}{1{\cdot}70 \times 0{\cdot}78} = 5700.$$

(*b*) *Second approximation.* The main correction to the first approximation comes from the new value of p (3·14) which will be evaluated for $y_0 = 0{\cdot}88$; the term $-z^2$ which was neglected in the expression for $w - c$ will also be roughly accounted for.

Put

$$w - c = y_0^2 - y^2, \quad y_0 = 0{\cdot}88. \tag{6·6}$$

In this approximation instead of assuming that (5·15)

$$w - c = w_0' z,$$

use is made of the expression (6·6); and a simple computation shows that in the present case

$$w - c = 1{\cdot}07 w_0' z. \tag{6·7}$$

This value of $w-c$ is inserted not only in the first term of ψ_2 (3·24), but also in the critical term; i.e. assume that

$$(w-c)\log(y_0-y) \approx 1{\cdot}07w_0'(y-y_0)\log(y_0-y), \tag{6·8}$$

and then transform $(y-y_0)\log(y_0-y)$ according to (3·23); this is not entirely correct, since the continuation of $z^2\log(-z)$ is not quite equal to the continuation of $z\log(-z)$ multiplied by z.

These computations lead to
$$\frac{w_0}{w_0''\epsilon} = -19{\cdot}8, \tag{6·9}$$
where
$$y_0 = 1-3/22{\cdot}8 = 0{\cdot}87, \tag{6·10}$$
which is close to the assumed value.

In the evaluation of α the quantity A_0 (5·9) was computed by making use of (6·6); whence finally

$$\alpha = 1{\cdot}0, \quad c = 0{\cdot}23, \tag{6·11}$$

and the critical Reynolds number is

$$R = 6700. \tag{6·12}$$

The coefficient of the term α^4 in ψ_2 was also evaluated; its contribution is small.

According to experimental evidence (Goldstein 1938, p. 321) laminar flow in a pipe remains stable for $R = 2\times10^4$ and even for $R = 5\times10^4$; this R is, however, evaluated from the expression

$$R = \frac{4mU_m}{\nu},$$

where m is the mean hydraulic radius (area of the section divided by its perimeter), and U_m is the mean flow. Reduced to the present definition of R, the critical number 5×10^4 becomes $R = 1{\cdot}9\times10^4$, which is of the same order of magnitude as (6·12).

According to Squire (1938), the critical R of a three-dimensional flow should be higher than for a two-dimensional one, i.e. higher than (6·12). The critical R is very sensitive to the value of ϵ, and it remains to be seen whether a more rigorous calculation will increase the value, or whether flow in a channel is less stable than in a pipe.

7. EQUATION OF MEAN MOTION

Consider the equation of motion in the direction of the x axis together with the equation of continuity

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y} = -\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\nabla^2 u, \quad \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y} = 0. \tag{7·1}$$

It is clear that in the present case

$$\overline{u\frac{\partial u}{\partial x}} = \frac{\partial}{\partial x}\frac{\overline{u^2}}{2} = 0, \tag{7·2}$$

where the bar above denotes the mean value; since

$$\overline{u\frac{\partial u}{\partial x}}+\overline{v\frac{\partial u}{\partial y}} = \overline{u\frac{\partial v}{\partial y}}+\overline{v\frac{\partial u}{\partial y}} = \frac{\partial}{\partial y}\overline{uv}, \tag{7·3}$$

then from (7·1) and (7·3) by integration

$$\left.\begin{aligned} \overline{uv} &= \frac{P}{\rho}y+\nu\frac{dU}{dy}, \\ P &= -\frac{\partial\overline{p}}{\partial x}, \end{aligned}\right\} \tag{7·4}$$

where

and U is the mean velocity.

Since

$$u = \frac{\partial\psi}{\partial y}, \quad v = -\frac{\partial\psi}{\partial x}, \tag{7·5}$$

it can be shown that

$$\overline{uv} = \sum_{\alpha} i\alpha(\psi_\alpha\tilde{\psi}'_\alpha-\tilde{\psi}_\alpha\psi'_\alpha), \tag{7·6}$$

where the sign ~ denotes a conjugate complex quantity and the differentiations are taken with respect to y.

Certain properties of the mean motion at the boundary can be immediately derived. Since ψ and ψ' vanish on the boundary it is easily deduced from (7·4) and (7·6) and by differentiation that

$$-\mu U' = Py, \quad -\mu U'' = P, \quad U''' = 0, \quad U^{\text{iv}} = 0, \tag{7·7}$$

also hold on the boundary; thus the flow close to the boundary is to a great extent of parabolic type, although with a higher pressure gradient than occurs in the case of laminar flow.

From the expressions (7·4) and (7·6) the mean motion close to the axis can be found; only one frequency is considered.

Put

$$\psi = d\psi_2, \tag{7·8}$$

where d is a constant and ψ_4 can be neglected near the axis; whence (3·7)

$$\psi = d\left\{(w-c)+\frac{i(w-c)}{\alpha R}\int_0^y \frac{w'''}{(w-c)^2}\,dy\right\}. \tag{7·9}$$

Substituting (7·9) in (7·6) it is found that (7·4)

$$\frac{2d^2}{RU_0}U''' = \frac{P}{\rho}y+\nu U', \tag{7·10}$$

whence integrating

$$U = U_0-\frac{Py^2}{2\mu}-B+B\cosh ky, \quad k = \frac{U_0}{d\sqrt{2}}, \tag{7·11}$$

where B is a constant of integration.

The turbulent complex stream function is

$$\psi = d(w-c)\,e^{i(\beta t-\alpha x)}, \tag{7·12}$$

the real stream function

$$\psi+\tilde{\psi} = 2d(w-c)\cos(\beta t-\alpha x), \tag{7·13}$$

the corresponding v velocity is

$$v = 2d(w-c)\,\alpha\sin(\beta t-\alpha x), \tag{7·14}$$

whence its amplitude round the axis is

$$2d(w-c)\,\alpha. \tag{7·15}$$

Comparing a corresponding experimental curve of the mean motion round the axis with (7·11) the constant d can be found, and from (7·15) the amplitude of v on the axis; the latter can also be found experimentally.

REFERENCES

Goldstein, S. 1936 *Proc. Camb. Phil. Soc.* **32**, 40.
Goldstein, S. 1938 *Modern Developments in Fluid Dynamics.* Camb. Univ. Press.
Heisenberg, W. 1924 *Ann. Phys., Lpz.*, **74**, 577.
Meksyn, D. *Proc. Lond. Math. Soc.*; in publication.
Noether, F. 1926 *Z. angew. Math. Mech.* **6**, 232.
Pekeris, Ch. 1936 *Proc. Camb. Phil. Soc.* **32**, 55.
Pekeris, Ch. 1938 *J. Aero. Sci.* p. 237.
Southwell, R. V. & Letitia Chitty 1930 *Phil. Trans.* A, **229**, 205.
Squire, H. B. 1933 *Proc. Roy. Soc.* A, **142**, 621.
Tietjens, O. 1925 *Z. angew. Math. Mech.* **5**, 200.
Tollmien, W. 1929 *Nachr. Ges. Wiss. Göttingen*, p. 21.

Experimental production of red and yellow sandstones from chalybeate water

By Lord Rayleigh, F.R.S.

Received 4 *June* 1945—*Read* 15 *November* 1945

This paper discusses the conditions for the formation of red sandstones, to some extent from the point of view of field geology, but mainly from that of laboratory experiment. Field observation made in the Tyne valley confirmed a suggestion made to me by Sir Robert Robertson that, in some cases at least, red sandstone may derive its iron from chalybeate water, which is a solution of ferrous bicarbonate. Starting with this idea, it was found by experiment that red or yellow sandstone could be produced according to the circumstances under which the chalybeate water deposited its iron. If the chalybeate water is evaporated on the surface of a solid body, whether a sand grain or the surface of a glass bulb, a red coating results. The same occurs when evaporation occurs in air from an undisturbed liquid surface, a red skin forming on the liquid. If, on the other hand, the liquid is agitated with air, as in a stream in the open, a light yellow precipitate forms throughout it. This when formed settles down as a yellow deposit on sand grains, twigs, or other solid surfaces, or on solid sandstone rocks.

The red and the yellow deposits both consist of hydrated ferric oxides. It is difficult to say from a chemical point of view what the essential difference of constitution between them is. The degree of hydration is not rigidly constant for different red preparations, and tends to be somewhat less for the red material than for the yellow.

The main point is that red sandstone has been produced which closely imitates the natural product. This has been done without using high temperatures and using only material (chalybeate water) of common occurrence in nature. The natural red grains were mounted side by side with the artificial ones and examined with the microscope using a low power and a dark ground illumination. While both samples showed a considerable range of intensity in the red colour of the individual grains, they were as a whole indistinguishable as to the colour and its mode of distribution.

It is believed that dehydration by volcanic heat or other agency, either of the sandstone itself or of its constituents, has nothing to do with the matter.

The colour of the rocks in red sandstone districts contrasts very strikingly with the yellow colour of the clays and sandstones in the home counties in the neighbourhood of London. In both cases the colour is, of course, due to iron, and the question presents itself: What are the essential physical conditions for the production of red sandstone rather than yellow? Guided by advice from the Geological Survey, I have searched the literature for any account of what the essential conditions are, but after consulting a large number of memoirs on the subject of the formation of red rocks, I was not able to find any account of laboratory experiments directed to this question. Preliminary tests were made to find at what temperatures yellow ferrugineous materials, e.g. London clay, brickearth, yellow flints, would begin to go red. They were not found to do so below about 200° C. Apart from volcanic action, such temperatures are not met with on the earth's surface. The record high temperature is said to be 56° C, recorded in Tripoli. As a modest contribution to the subject I wished to discover some method by which red sandstone could be produced in the laboratory without making use of temperatures higher than those met with in the tropics; to use only materials which nature is known to provide; and at the same time to distinguish as clearly as might be between the conditions which would give yellow sandstone and red sandstone respectively. If such a method were found it might at least afford a possible answer to the problem set by nature.

My attention was first turned to the solution of ferrous bicarbonate by Sir Robert Robertson, who kindly interested himself in the matter and gave me various useful references. This solution, analogous chemically to calcium bicarbonate, occurs in chalybeate springs. The presence of such springs is recognized by the bright yellow precipitate of ferric hydrate which separates out as they emerge from the earth and spread out into the open. They are numerous in Northumberland in the valley of the Tyne, or rather of its tributary known as the Corburn, which flows into it at Corbridge. There is probably nothing very notable about these particular chalybeate springs, but circumstances made them easy of access to me. The springs I have examined are at Halton Castle, and along Aydon Dene, just below Aydon Castle.

The local (carboniferous) sandstone, such as that from the quarry near Beaufront Castle, while occasionally almost colourless, is usually yellow or brown or dark brown. Sometimes the coloured layer is rather superficial. Sometimes it is vein-like, following what has been a crack in the rocks. The chalybeate water issues from the ground and stains the sandstone in this quarry, and, after examining it, I could scarcely doubt that the sandstone of the district generally, when coloured yellow or brown, derived its colour from chalybeate water.

I have said that the stone is yellow or brown, but occasionally, as in Halton chapel, red pieces are seen, which superficially, at any rate, resemble the colour and appearance of old red sandstone, or other red sandstone formations. I also collected a hand specimen in which the ordinary yellowish brown colour is seen to merge into red. This was from a heap of broken stone by the roadside. So far as I have been able to learn, Halton chapel was probably built of stone from a quarry now overgrown and almost forgotten at Carr Hill close by. I was able to find one deep red-coloured piece in the sandstone there, but its relation to other pieces, and to the undisturbed rock, could not be clearly traced. The red is rather rare and apparently quite sporadic, but its occasional presence along with the brown seems to make it very probable that its colour too is due to chalybeate water, and I am tempted to generalize and to lean to the view that the same is true of all red sandstone formations.

These field observations probably contain little that is new to geologists. Any value they may have depends upon the laboratory investigation which follows. The problem now became this—to specify methods whereby sand could be coloured red or yellow at pleasure, using chalybeate water as a source of iron, and avoiding the use of temperatures higher than about 50° C. The cementing of the sand into sandstone might be taken for granted.

The work was continued in my laboratory in Essex where no natural chalybeate water was conveniently available and artificial chalybeate water was substituted. It was prepared by the method of G. T. Moody (1906), distilled water being allowed to stand over a mass of fine iron wire, and carbon dioxide passed in for a day or preferably more. Moody obtained a solution containing 0·225 g. FeO/l. in 20 hr. and 1·39 g. in 26 days. The natural chalybeate water of Langenswalbach contains, according to Fresenius, 0·376 g. FeO/l., and is thus of about the same concentration as the artificial product made in a day or two. The latter was drawn off by a pipette or siphon, avoiding exposure to air as far as possible.

If some of the chalybeate solution is drawn off into a test-tube and carefully examined from time to time without disturbance, two things may be noticed. First a canary yellow precipitate forms in the body of the liquid, and subsides into a yellow layer at the bottom of the tube. Secondly, a skin of red colour forms on the surface, and progressively increases in thickness from day to day. To see these simultaneously for comparison, it is desirable to immerse the lower part of the test-tube in water, which mitigates the reflexion from its outer surface, and makes it easier to see the layer of yellow precipitate lying at the bottom. The true surface of the chalybeate water which is clear of the external immersion vessel can be seen at the same time, and the colour contrast of the red skin and the yellow precipitate is very striking.

If a spiral of clean iron wire is sprung into the tube, extending over most of the liquid layer, but leaving it free for observation at top and bottom, the yellow precipitate is not formed in the body of the liquid, nor, of course, is there any eventual formation of a yellow layer at the bottom. Hydrogen is given off from the iron, and its function is presumably to maintain reducing conditions except at the top, where the air has access. The red skin still forms as before. Assuming, as it seems necessary to do, that the yellow material is in some way distinct in composition or molecular structure from the red, we keep the red in this way free from the yellow. Since, in preparing chalybeate water, the liquid is standing on iron wire, formation of the yellow substance is not very noticeable, even if the current of carbon dioxide is discontinued. In the latter event, however, the red skin does continue to form.

If we wish to favour the formation of the yellow precipitate rather than the red skin, then the interior of the liquid should be brought into contact with air as much as possible, in other words, air should be bubbled in. Under these circumstances all the chalybeate water is soon oxidized to the yellow material, and there is no opportunity for the formation of the red skin before the dissolved ferrous carbonate is exhausted. Moderate heating, which is all that was tried, makes no difference. The same thing happens in nature when a chalybeate spring emerges into the open and the water flows away with appreciable turbulent movement. The surface gets broken up, and there is no formation of a red skin. An abundant yellow precipitate of ochre is formed.

(In some cases, however, as in a seepage of weak chalybeate water in my own garden, a thin skin does form, but not enough to show its own red colour. It gives a silvery reflexion, which is clearly the 1st order tint of Newton's scale, and it may be seen passing into tints of higher order. To observe this, leave some ordinary tap water, which of course contains carbon dioxide, standing in a cast iron vessel for a day or two, and watch the reflexion from the surface.)

It appears probable that the oxidation of ferrous iron rather than the loss of carbon dioxide is the cause of precipitation. The evolution of carbon dioxide without oxidation would presumably lead to the precipitation of ferrous carbonate, whereas the material actually precipitated is not carbonate but hydrate.

Why there should be the two different kinds of precipitate (red on the surface and yellow in the volume) there is no direct evidence to show. It seems plausible to suggest that the surface effect is produced by contact with gaseous oxygen, and the volume effect by dissolved oxygen.

When the skin is thick enough to be detached it shows its own characteristic absorption colour, which varies according to the thickness from a pale yellow to a deep red. The thickness of such a skin may be estimated under the microscope viewing a fragment edgewise, but this method does not lend itself to the use of a high power and only a rough estimate could be made. A film of about 16μ thickness gave a dark red. A film estimated at about 2μ in thickness gave an orange colour while the average sand grain in a very dark specimen of old red sandstone only gave a yellow. The grains of a red sandstone have only a superficial layer of colouring matter, and if this is washed away with acid, colourless quartz remains, as in ordinary sand. We see then that the coloured grains of this sand are scarcely as red, when examined by transmission, as a skin 2μ thick on chalybeate water. It appears therefore that the material of this skin is quite red enough to satisfy the requirements.

The colour of red sandstone is, or was, often regarded as due to haematite. The colour of the red skin as formed from chalybeate water is not due to the anhydrous peroxide: this may be proved most simply as follows. If we take a glass bead coloured with the red skin in the way described below (figure 1), and heat it in a small porcelain crucible, the bead becomes *very* much darker, a layer initially orange becoming dark red, and almost opaque. It has in fact been dehydrated and converted to haematite by heating, and was not haematite initially.

This experiment of the permanent deepening of colour by heat may also be made on the sand grains coloured by chalybeate water, and on the similar red grains of the old red sandstone. However, the coloured glass beads show it much more satisfactorily, on account presumably of their smoother surface and better transparency.

The transmission of light by the chalybeate skin formed on the liquid surface may be compared directly with the transmission by haematite. Only the thinnest flakes of micaceous haematite are observed to be at all transparent. Those that are so are about 2μ in thickness and show only the deepest red. A chalybeate skin of this thickness shows only an orange colour; it requires a thickness of some 16μ to show a deep red like the haematite flake.

Other considerations also tell against the view that the colouring matter of the red skin is haematite. It has not the characteristic full red streak (i.e. colour of powder) which is given by haematite, but inclines much more to orange. Further, it is hydrated to a considerable extent. The loss on ignition is some 17 % (see below). The loss on ignition by Cumberland haematite is almost nil (0·2 % was found).

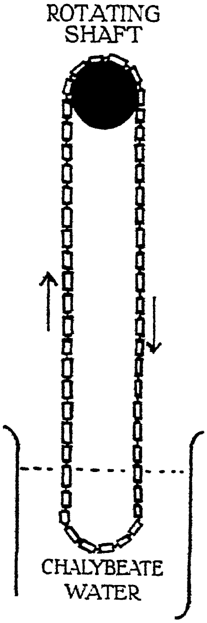
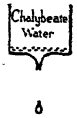


FIGURE 1

The question naturally arises whether the skin which forms on chalybeate water is essentially different from the yellow precipitate which forms when air is passed in, for it has been seen that the skin too is yellow in small thicknesses. I should be glad to be able to simplify the subject by regarding these as identical. The degree of hydration does not appear to be greatly different. The chief difficulty arises in comparing the body colour and the streak. These are both much redder in the case of the skin than in the case of the yellow precipitate formed by bubbling in air. The skin ground up gives an orange streak and the sand grains in bulk well coated with it give the full red colour of red sandstone. The precipitate supplied in any quantity by chalybeate streams is pale yellow, or, on drying, brownish yellow, and in no circumstances red, whether alone or coated on sand, and it gives a yellow streak.

It seemed probable that the formation of the red skin is quite independent of the depth of the liquid on which it is formed, and if so it should be capable of forming on the layer of liquid which coats any wet solid body. Further, if the wetting and drying are repeated, the layer should be progressively thickened, just as if it had been formed by prolonged evaporation from an ordinary free liquid surface.

To test these ideas, some beads formed of short lengths of glass tube were threaded on a string to form a kind of necklace (figure 1). This necklace was hung on the pulley of an electric motor with slow worm gear, so that each bead on the necklace moved with a linear velocity of 4 cm./sec. The bottom of the necklace dipped in a vessel of chalybeate water, and the free part of the length from the liquid to the pulley (about 1 m.) and on from the pulley to the liquid was exposed to the radiation from an ordinary domestic electric stove, which dried it off. After this it was again wetted and dried and so on for as long as desired. The beads became yellow in say 15 min., then orange, then red by the thickening of the layer.



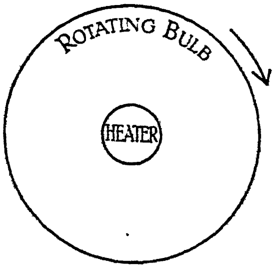


FIGURE 2

To collect the red material for analysis, a slightly modified arrangement was used (figure 2). A large glass flask 21 cm. in diameter was slowly rotated about the neck as axis with a circumferential velocity of about 6 cm./sec. It was kept hot by an electrical resistance heater inside. The chalybeate water was caused to drip upon it, at a regular rate, so that an equatorial belt was kept covered with a film of the liquid and was constantly drying off and being replaced by fresh. In this way a solid deposit was formed on the equatorial belt, first yellow, then red, and of firm texture. After a run of some hours, it was scraped off with a safety razor blade and collected. 50 mg. or so could be made in a day.*

These observations suggest a way in which either yellow or red sandstone can be produced at pleasure from chalybeate water.

To produce yellow sandstone, the sand grains are well under the surface of the liquid, which is evaporated, e.g. by the rays from an infra-red lamp,† which serves for this purpose as the equivalent of tropical sunshine. The chalybeate water yields a yellow precipitate which coats the sand grains. The evaporation is, however, more quickly carried out on a water bath, and the result is the same.

---

* In setting up a new arrangement of this kind, I should try a large glass cylinder in preference to the spherical flask.

† As supplied for medical purposes in the treatment of rheumatism and the like. It is merely an electrically heated source worked at a not very high temperature at the focus of a deep concave mirror.

To produce a red sandstone the sand grains must be only moistened with the chalybeate water, their moistened surfaces being exposed to the air, and to the heat rays from the 'infra-red lamp'. The skin, equivalent to the red skin which forms on the surface of chalybeate water allowed to stand in the open, will then form on the surface of the wet sand grains, and the process must be allowed to continue until the individual grains, mounted in balsam and viewed under the microscope, have the same depth of transmitted colour as the grains of a well coloured specimen of old red sandstone. When this has been done, such grains dried and compared with the natural sandstone grains will be found to be closely similar to them, whether viewed by transmitted or by scattered light.

The iron content of the red sandstone which it is desired to imitate can readily be determined by reducing the solution to the ferrous condition in the usual way, and titration with potassium permanganate.

The chalybeate water which is used requires no reduction, and can also be titrated with permanganate. It was thus found that the selected sample of old red sandstone contained 0·365 % of iron, and that the chalybeate water, as I prepared it, contained 0·138 g. Fe/l. Thus, to make an artificial red sandstone comparable with the natural, we should expect that 26·4 c.c. of the chalybeate water would be required per gram of sand.

Some such quantity of chalybeate water must in fact be evaporated from the wetted surface of the sand grains to get the desired effect. The sand must not be too wet, for in that case the water covers the grains, sealing them off, and hindering the necessary evaporation from the surface of each wet grain. I have attempted to maintain the right degrees of wetness automatically. This could doubtless be done, but it is not quite easy. If we introduce the water locally, dropping from a fine glass jet, then the jet gets plugged by a slight precipitate from the chalybeate water. This can be overcome by the use of a suitably designed filter. A more serious difficulty is that the sand gets too wet near the jet. Further out, as the chalybeate water soaks into the sand and dries up, we come to a region where there is no chalybeate water and no action. I have succeeded best by placing the sand (100 g. say) in a shallow vessel made by laying an iron hoop 10 cm. in diameter and 1 cm. high on a glass plate. Chalybeate water is added from a pipette until the sand is just wet throughout, any slight excess draining away between the hoop and the glass plate. It is then allowed to dry up completely, under the heat rays, and only then is more chalybeate water added. The additions were at about half hourly intervals.

An alternative method of keeping the sand layer moist but not flooded is to allow the chalybeate water to rise by capillarity through a sand column conveying it to a sand layer in a metal tray. The tray is warmed to 50° C in an oven. As the evaporation proceeds the sand layer becomes red where the supply of chalybeate is most abundant, and yellow further out. This arrangement perhaps represents pretty closely the way in which a sand surface is supplied with chalybeate water under natural conditions. Evaporation would occur from the exposed surface heated by the sun which I have imitated by means of the infra-red lamp, or by the oven (figure 3) maintained at 50° C.

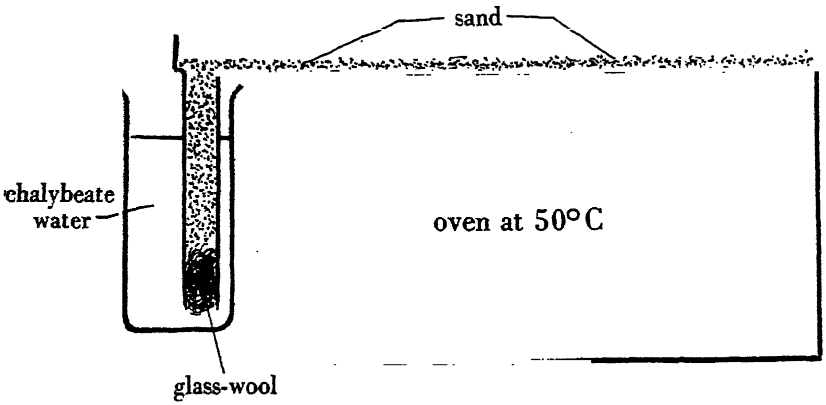


FIGURE 3

Nature, working on a large scale, more easily maintains the right degree of moisture. But it may be remarked that where large exposures of red sandstone are examined, it is often, if not always, found that the depth of colour varies from place to place, often from white up to a deep red. This suggests that accidental and local circumstances determine whether and for how long the best conditions are maintained.

In preparing the red sand experimentally, I took a specimen of old red sandstone from Dartmouth as a model of what to aim at. This was the reddest that could be easily found there. It was broken down into sand with little loss of the iron colouring matter and some of the grains mounted in balsam. These were viewed with a lens or low power microscope and served as a standard for the depth of colour. They were for the most part just red, rather than yellow, by transmitted light. Some sand, sifted so as to pass a sieve of 30 to the inch and to be stopped by a sieve of 60 to the inch was cleaned of all iron with hydrochloric acid, washed and dried, so as to start fair. 100 g. was taken and it was drenched with 20 c.c. of chalybeate water which was about as much as it could absorb. It was allowed to dry under the 'infra-red lamp' and then drenched again, and so on, until about 2 l. of the chalybeate water had been added. At this point it was found that a few grains of the sand, mounted in balsam and examined under a lens by transmitted light, had about the same average depth of colour as the grain from the natural red sandstone. The quantity of chalybeate water required for this (20 c.c./g. of sand), though in accordance with the above calculation, was much more than I had originally imagined, but, after all, it should have been obvious that if we are to imitate the natural sandstone of about the same grain size we must add enough iron to bring the iron content of the whole up to the same value.

When the individual grain viewed by transmitted light had been brought up to somewhat the same depth of yellow or red coloration as the grains of the natural red sandstone, then the aggregate viewed by reflected or diffused light agreed closely with it in general appearance. To verify this satisfactorily it is necessary to view the two side by side under similar conditions. I have found that the natural red sandstone under some conditions of lighting had an appearance very different from the mental idealization I had formed of it, and it may be well to put others on their guard on this point.

It was found, comparatively late in the investigation, that the individual red grains, mounted as usual in balsam, can be far better examined by dark ground illumination than by ordinary transmission. No special apparatus is required except a black central stop of 10 mm. diameter over the condenser of the microscope. A $\frac{2}{3}$ in. power is suitable. Narrow streaks and patches of intense blood red colour are then seen in the individual grains. The appearance is characteristic and striking, and the colour and its peculiar distribution are exactly comparable in the natural old red sandstone and in the artificial imitation. These may be mounted so as to be viewed at one coup d'œil under the microscope with a partition across the field of view. They were shown in this way when the present paper was read, and the comparison of natural and artificial was, I believe, found convincing. Unfortunately this mode of illumination does not give enough light for screen projection before an audience.

The grains of yellow or brown sandstone, whether natural or artificial, examined under dark ground illumination show the same distribution of colouring matter as the red. It is obvious that this distribution is due to the holes and crevices in the grains, which afford a lodgement for the ferrugineous colouring matter, whether in nature or in the laboratory.

It may, I think, be claimed that the imitation of natural red sandstone is completely successful.

It will be asked, what is the chemical nature of the yellow precipitate which forms when air is bubbled into chalybeate water, and wherein does the red skin which forms on a surface differ from it?

It is to be noticed that no materials are present except iron, distilled water, and carbon dioxide, so that there is comparatively little room for chemical complications. Further, on dissolving the precipitate or the red skin in hydrochloric acid, and testing with ferricyanide, no blue precipitate is obtained. Ferrous compounds may therefore be ruled out. The red and yellow products dissolve in acid without effervescence, so that there is no question of carbonates. Moreover, no such substance as ferric carbonate is known to exist. There appear, therefore, to be no possibilities left except the various ferric hydrates and it remains to find the loss of weight on ignition in order to determine the percentage of combined water. The materials have been ground up fine in an agate mortar, and left standing over sulphuric acid, for many hours in a high vacuum. Drying by heat was avoided, because it was not desired to go into the question of the possible loss of combined water in this operation. Some doubts were expressed at the reading of the present paper as to whether free water was effectually got rid of. Personally I do not share these doubts. The production of a high vacuum is conclusive evidence. Anyone with experience of evacuating X-ray tubes and the like knows the futility of attempting it until every trace of water is gone.

Treating the yellow precipitate obtained by blowing air through artificial chalybeate water in this way, the following determinations were made:

percentage loss on ignition	mean
19·0, 21·6, 18·9	19·8

For comparison with the above, some of the precipitate from the natural chalybeate water at Aydon Dene near Corbridge, Northumberland was examined in the same way.

percentage loss on ignition	22·7, 19·3, 21·0
mean	21·0

in fair agreement with the results for the artificial preparation.

The red skin formed on the surface of undisturbed chalybeate water was ground to a fine powder in an agate mortar, and dried in vacuo as in the previous case.

specimen number	percentage loss on ignition	mean
1	17·2	17·2
2	17·8, 18·0, 17·3	17·7

These values are lower than those obtained for the yellow precipitated material, though the difference is scarcely large enough for it to be possible to build very much upon it.

With regard to the red material, obtained as above by evaporation of chalybeate water on the surface of a glass flask internally heated, since the chalybeate water never boiled on reaching the glass surface, it may be concluded that the evaporation occurred at a temperature less than 100° C.

About 20 mg. of the product were placed in a small platinum basin weighing less than 2 g., and dried in vacuo. The loss of weight on ignition was determined, using a delicate assay balance. It was found in various experiments to have the following values:

specimen number	percentage loss on ignition	mean
1	11·5, 12·4, 13·4	12·4
2	14·1, 15·3	14·7

These values are decidedly lower than those obtained with the ground-up red skin, a result which may perhaps be attributed to the partial dehydration resulting from the process of drying off by heat on the rotating bulb. In this method of preparing the material, heating cannot well be avoided, since without it the process would be too slow.

It may be mentioned for comparison that the composition $Fe_2O_32H_2O$ corresponds to 18·4 % water, and $Fe_2O_3H_2O$ to 10·1 % water. The determinations of combined water were carefully carried out. The results, though fairly concordant for a given specimen, show considerable variation from one specimen to another. The conditions of drying on the surface of the rotating bulb could not be held rigidly constant, nor were the conditions of temperature well defined, since the heated surface of the bulb was in the open. There was, however, no ebullition when the drops fell on the bulb, and the temperature is believed to have been well below the boiling point. Upon the whole the statement seems justified that the red material is somewhat, though not necessarily very much, less hydrated than the yellow. It is at any rate very far indeed from being anhydrous.

The present work, it is hoped, shows clearly enough that chalybeate water can provide either yellow or red sandstone according to circumstances. If the chalybeate water is churned up with air, it deposits yellow ochre, which will coat the surface of any solid, such as leaves, twigs, solid rocks or sand grains. This is the origin of yellow sandstone.

If the chalybeate water evaporates from its own undisturbed surface or on a wet glass surface, or on a wet sand grain, it leaves a film yellow in small thicknesses and red in greater ones. In this way red sandstone can be produced. Contrary to what has often been supposed, the red material, like the yellow, is hydrated, and to a comparable extent. This is the origin of red sandstone, which has nothing to do with the action of heat as, for example, in the baking of yellow clay into bricks. All this seems to be well supported by the experiments described.

Why red sandstones should be characteristic of particular geological epochs, and yellow ones of others, is a question which I do not attempt to answer. I have limited myself to showing how they can be produced *now*, using materials which commonly occur at the earth's surface under ordinary temperature conditions. It is hoped that this will be a contribution to the subject, and may lead to further progress in answering the mysterious questions which the geology of red sandstones presents to us.

APPENDIX

Further experiments on the formation of red iron oxide at ordinary temperature

It was natural to try the use of dry ferrous carbonate instead of chalybeate water (which is a solution of ferrous carbonate), the more so that pseudomorphs of haematite after ferrous carbonate are known to occur. Ferrous carbonate occurs as chalybite and also under a somewhat different form as spathic iron ore, but these substances do not show any visible signs of oxidation when finely powdered and heated in air to 100° C for 24 hr. They were therefore left out of account in further experimenting, though without prejudice to the question of what might occur in very long periods.

On the other hand, freshly precipitated moist ferrous carbonate oxidizes quickly in the air at ordinary temperature, giving yellow or brown ferric hydrate, just as chalybeate water does when exposed or allowed to dry in an open vessel.

Ferrous carbonate may, however, be precipitated in the absence of air, and dried in vacuo over sulphuric acid. It then oxidizes in air at about 50° C (and possibly also at room temperature) to a red hydrated oxide.

REFERENCE

Moody, G. T. 1906 *J. Chem. Soc.* 89, 720–730.

The probability of encounters between gas molecules

BY LEWIS F. RICHARDSON, F.R.S.

(*Received* 25 *June* 1945)

Two unusual measures of the occurrence of binary, ternary, ... encounters are given, each by an appropriate line of reasoning. Both indicate deviations from the law of Guldberg and Waage at high densities.

First, the close analogy between gases and emulsions, which was established by the experiments of Perrin, is brought into relation with experiments by W. S. Gosset on the distribution of the square cells of a haemacytometer as to the number of particles which they contain. This analogy indicates that, in a cubic centimetre containing n molecules, the number of *simultaneously existing* encounters of 2 or 3 or 4 molecules in a bunch will respectively be

$$\tfrac{1}{2}n^2\omega_1 e^{-n\omega_1}, \quad \tfrac{1}{6}n^3\omega_1^2 e^{-n\omega_1}, \quad \tfrac{1}{24}n^4\omega_1^3 e^{-n\omega_1},$$

where ω_1 is a constant of the order of 10^{-22} cm.^{3}.

Secondly, the problem is re-examined by a statistical argument after the manner of Bernoulli and Poisson, and various improvements are inserted. If there are n_1 molecules of one species and n_2 of another in a cubic centimetre then, during 1 sec. of the observer's time, the *aggregate duration* of binary encounters between unlike molecules, all of diameter $\sigma = 3\times 10^{-8}$ cm., is found to be

$$n_1 n_2 \Omega \exp\{-0{\cdot}79(n_1+n_2)\,\Omega\},$$

where

$$\Omega = \omega/\{1-(n_1+n_2)\,\sigma^3 2^{-\frac{1}{2}}\},$$

and ω is of the order of 7×10^{-23} cm.^{3}.

To specify a binary encounter it does not suffice to say that two molecules must meet; for one must also say that an enormous number of other molecules must keep away from the pair. The latter requirement, although obvious when mentioned, and well known in statistical theory, appears to have escaped notice in previous theories of collisions in a gas. The exponential factors in the above formulae express the probability that the extraneous molecules will keep away.

The treatment emphasizes those statistical proprieties, but otherwise proceeds by easy approximations; so that the better determination of ω, for particular reactions, is left to experiment.

FIRST THEORY

This is a variation on W. S. Gosset's theme concerning the haemacytometer (1907). In the haemacytometer a thin layer of emulsion, of uniform thickness, lies on a glass plate ruled into many small equal squares. Gosset (alias 'Student') counted the number of squares which contained respectively 0, 1, 2, 3, ..., r, particles and found it to be

$$Ne^{-\mu}\mu^r/r!, \tag{1}$$

where N is the whole number of squares and μ is the average number of particles in a square. The formula is Poisson's law of rare statistical events.

To apply these ideas to a gas, let a vessel in the form of a centimetre cube be subdivided in imagination by planes parallel to its walls so as to cut it into a large number of cubical cells each of volume ω_1. The vessel thus becomes a three-dimensional analogue of the haemacytometer. Let the vessel contain n molecules, and, for simplicity, let each of these be regarded as a sphere of diameter

$$\sigma = 3\times 10^{-8}\,\text{cm.} \tag{2}$$

It is essential to the theory that ω_1 should be chosen judiciously, neither too small nor too large. There must be room in a cell for the pair, or triplet, of molecules that are to take part in the chemical reaction. Yet, when they are together inside a cell, they must be so close together that conditions for chemical reaction are favourable. In other words it is proposed to regard the simultaneous presence of the reagent molecules inside the cell as constituting an encounter. Opinions will differ as to the proper size of the cell. But, for the purpose of estimating orders of magnitude, take the edge of the cube to be twice the molecular diameter, so that

$$\omega_1 = 8\sigma^3 = 2{\cdot}16\times 10^{-22}\,\text{cm.}^3. \tag{3}$$

In view of the close similarity between emulsions and gases, made manifest by the work of Perrin, some readers may be willing to jump directly by analogy from Gosset's observations to the assertion that the number of cells which each contain r molecules, and no more, is given by Poisson's law (1), in which N, being the total number of cells, is

$$N = 1/\omega_1, \tag{4}$$

and μ, being the average number of molecules per cell, is

$$\mu = n\omega_1. \tag{5}$$

With these substitutions, the number of cells which each contain just r molecules is

$$n^r\omega_1^{r-1}\,e^{-n\omega_1}/r!. \tag{6}$$

That is to say $e^{-n\omega_1}/\omega_1$ cells are empty, $ne^{-n\omega_1}$ cells each contain a solitary molecule, $\frac{1}{2}n^2\omega_1 e^{-n\omega_1}$ cells each contain a binary encounter, $\frac{1}{6}n^3\omega_1^2 e^{-n\omega_1}$ cells each contain a ternary encounter, and so on. One must not, however, continue much further, lest the cell should become choked with molecules. A correction for the gradual filling of the cell is made in the second theory. The customary measure of encounters is the number that begin during a second. A different measure has now been obtained, namely, the number that simultaneously exist. Provided that each encounter has a finite duration, the present measure is quite as intelligible as the customary one.

In particular, it follows from (6) that

$$\frac{\text{Ternary encounters existing}}{\text{Binary encounters existing}} = \frac{n\omega_1}{3}. \tag{7}$$

At N.T.P. it is well known that

$$n = 2 \cdot 7 \times 10^{19} \text{ cm.}^{-3}. \tag{8}$$

In those circumstances, it follows from (3) and (8) that $\frac{1}{3}n\omega_1 = 0 \cdot 002$. That is merely a rough estimate of the order of magnitude; but it resembles Bodenstein's estimate, namely, 0·001, obtained by a different theory quoted by Hinshelwood (1940, p. 142). Ternary encounters will be reconsidered in a later section.

The number of binary encounters existing simultaneously in a cubic centimetre is $\frac{1}{2}n^2\omega_1 e^{-n\omega_1}$. In this the factor n^2 is well known, and is in accordance with the law of Guldberg and Waage. But the factor $e^{-n\omega_1}$ may be novel, and represents a deviation from that law. The effect on bimolecular chemical reactions cannot be stated clearly until a mixture of molecular species is considered. But, as a preliminary, it may be noted that $n^2e^{-n\omega_1}$ increases less rapidly than n^2 and attains a maximum at $n\omega_1 = 2$; that is to say, when there are, on the average, two molecules in a cell.

The theory which has just been stated is an easy approach to the subject, but is not convenient for further developments, chiefly because cubical cells do not fit neatly round spherical molecules. So let us break off here, abandon ω_1, and make a fresh beginning from first principles.

SECOND THEORY

Poisson's law of rare statistical events is derived from the assumption that there are very many occasions and, at each, an almost zero probability of the event occurring. There is nothing in the deduction of Poisson's law (e.g. Sheppard 1929; or Aitken 1942) to say whether the occasions are many in space, many in time, or many in each of them. In the first theory about 10^{22} cubical cells were considered at the same instant. Now consider one cell throughout a long interval of time, which can, if desired, be supposed to consist of many successive occasions. An advantage of this change in procedure is that a single cell may be made of a shape to fit a molecule; for the cell need not be fitted to any neighbouring cells.

Jeans (1916, p. 371), when making an estimate of the duration of a binary collision, considered that: 'For an average molecule the effective diameter will be about 3×10^{-8} cms. (see p. 341), so that this may be supposed to be the distance apart of the centres of the two molecules at their closest approach. The force between them may, taking an average value, be supposed to fall off as the eighth power of the distance, and so will, at a distance of $3{\cdot}5 \times 10^{-8}$ cms., be equal to $\cdot 291$ times that at the distance of closest approach. The acute stage of the collisions may then, for a rough calculation, be supposed to extend from a distance $3{\cdot}5 \times 10^{-8}$ cms. to a distance 3×10^{-8} cms.' In order to reason clearly it is desirable to imagine that individual molecules can somehow be marked, named, and followed. Thus let A_1, B_1, C_1, ... be centres of particular molecules of the first species, and A_2, B_2, C_2, ... be centres of particular molecules of the second species. Let the single cell now be the space between two spheres having their common centre at A_1 and the radii 3×10^{-8} and $3{\cdot}5 \times 10^{-8}$ cm. chosen by Jeans. This cell, alias shell, surrounds a particular molecule and wanders everywhere with it. The volume of the cell is

$$2{\cdot}46\sigma^3 = 6{\cdot}65 \times 10^{-23}\,\text{cm.}^3 = \omega, \quad \text{say.} \tag{9}$$

That is smaller than the volume ω_1 of the cell of the first theory. In fact

$$\omega = 0{\cdot}308\omega_1. \tag{10}$$

There are inevitably differences of opinion here. Two ways out of the ambiguity will be suggested in the final section.

Let the measure of the probability of an event be the fraction of the observer's time during which the event occurs. In that sense the probability that the molecular centre A_2 is inside the spherical shell may reasonably be taken to be the ratio of the volume ω of the shell to the whole volume, available to molecular centres, in the vessel. At low densities the whole available volume is practically 1 cm.³. But the first theory showed a factor which becomes interesting at high densities. It is therefore desirable in this second theory to attend to the restriction of the available free space. The number of molecules, each a sphere of diameter σ, which could go into the vessel if they were closely packed would be $2^{\frac{1}{2}}\sigma^{-3}$ (W. L. Bragg 1933, pp. 144–5). Therefore

$$1-(n_1+n_2)\sigma^3 2^{-\frac{1}{2}} \tag{11}$$

is a reasonable measure of the whole volume available to molecular centres; for this expression is valid if the vessel is either quite empty or quite full, and is graduated between these extremes in a common-sense manner. If greater accuracy were required, it could probably be obtained from experiments on compression (Bridgman 1931, pp. 107–14). The probability that A_2 is in the cell is therefore

$$\frac{\omega}{1-(n_1+n_2)\sigma^3 2^{-\frac{1}{2}}} = \Omega, \quad \text{say.} \tag{12}$$

When A_2 is in the spherical shell there is less room in it for any other molecular centre. From the diagram, which shows the mean position of A_2 in the shell, it is plain that the excluded volume is on the average about that cut from the shell by a cone of semivertical angle 55°. This excludes a fraction

$$\tfrac{1}{2}(1-\cos 55^\circ) = 0{\cdot}213 \tag{13}$$

of the shell, and so leaves a volume, available for the centre B of a third molecule, equal to $0{\cdot}787\omega$. The probability that B will be in the shell is accordingly, by (12), $0{\cdot}79\Omega$; and the probability that B will be outside the shell is

$$1-0{\cdot}79\Omega. \tag{14}$$

Next assume that the events

'A_2 inside' and 'B outside', 'C outside', 'D outside', ... are statistically independent, alias uncorrelated. (15)

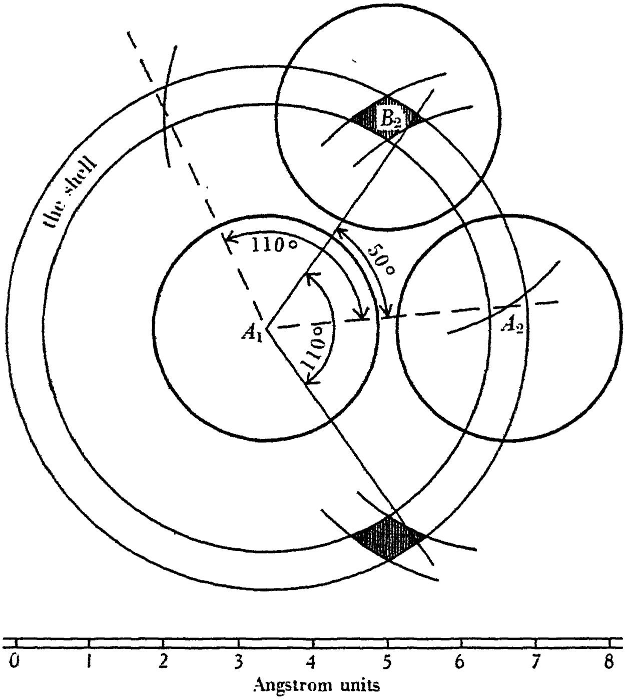


FIGURE 1

Jeans (1916, pp. 56–7) commented on the difficulty that his predecessors had experienced in the search for a strict dynamical proof of the absence of a different correlation in a gas. No proof of (15) will be attempted here. But (15) is a common-sense notion derived from practical experience in mixing powders, emulsions, or solutions. Here (15) is taken as part of the *specification of an abstract chaos*. It seems very likely that a gas closely resembles a chaos of that specification; but the final decision must be left to experiment.

There are n_1+n_2-2 molecular centres which must be kept out of the cell, in order that the encounter may be binary. The probability that they will all keep out is

$$(1-0{\cdot}79\Omega)^{n_1+n_2-2}, \quad \text{or practically} \quad (1-0{\cdot}79\Omega)^{n_1+n_2}. \tag{16}$$

This is the requirement, and the factor, which has been ignored in standard treatises on gases, and to which therefore attention is now particularly directed. Unless the molecules are very nearly close-packed, the quantity Ω is so near to zero that, with excellent approximation,

$$(1-0{\cdot}79\Omega)^{n_1+n_2} = \exp\{-0{\cdot}79(n_1+n_2)\Omega\}. \tag{17}$$

So this factor corresponds, apart from improvements, to the factor $e^{-n\omega_1}$ in the first theory.

Thus the probability that the molecule whose centre is A_1 is in encounter with the molecule whose centre is A_2, and that all other molecules are keeping away is

$$\Omega \exp\{-0{\cdot}79(n_1+n_2)\Omega\}. \tag{18}$$

Probability is here reckoned as a fraction of the observer's time.

But there are n_1n_2 pairs like A_1 and A_2. As it has been specified that there are only two molecules in the bunch, the presence of any one of these n_1n_2 pairs excludes the presence of all the other pairs. The aggregate time of occurrence of them all is therefore the sum of the aggregate times for each separately. That is to say, the aggregate duration of binary encounters between unspecified molecules in a cubic centimetre is

$$n_1n_2\Omega \exp\{-0{\cdot}79(n_1+n_2)\Omega\} \tag{19}$$

during 1 sec. of the observer's time. This last aggregate duration (19) can far exceed the observer's time, and is thus not a probability; but aggregate duration shares with probability the property of combining by addition when the events are mutually exclusive.

Expression (19) resembles $\frac{1}{2}n^2\omega_1 e^{-n\omega_1}$, which the first theory gave for the number of coexisting binary encounters in a cm.^{3}, but with the following modifications: (i) n_1n_2 replaces $\frac{1}{2}(n_1+n_2)^2$, because $\frac{1}{2}n_1^2$ and $\frac{1}{2}n_2^2$, which would appear in the formulae for encounters between molecules of like species, are being ignored; (ii) ω, based on Jeans's opinion, replaces the preliminary ω_1; (iii) the 0·79 represents the filling up of the cell; and (iv) the replacement of ω by Ω represents the filling up of the vessel.

If allowance be made for such modifications, it appears that:

> *The number of coexisting binary encounters is equal to the aggregate duration of all the binary encounters during* 1 *sec. of the observer's time.* (20)

([*Note inserted* 18 *July* 1945.] The italicized statement (20) is the main connexion between Theories I and II. But its purport becomes clearer when we temporarily ignore both these theories and instead proceed thus. In any vessel, let x be the number of binary encounters co-existing at the instant t, and let g be the aggregate duration of all the binary encounters during the interval $0 \leqslant t \leqslant 1$. Then, in any sufficiently brief time-element δt, the aggregate duration of all the binary encounters approximates to $x\,\delta t$. So

$$g = \int_0^1 x\,dt \tag{20A}$$

This (20A) is a strict result, not in the least qualified by probability. To pass from (20A) to (20), we need only assume that, because of the chaos, x is practically constant in time. For then (20A) becomes $g = x$, which is (20). Having finished this explanation, let us now return to Theory II.)

As the density tends to zero, Ω tends to a constant ω, and $\exp\{-0{\cdot}79(n_1+n_2)\Omega\}$ tends to unity, leaving in (19) only $n_1 n_2 \omega$. The factor which represents the deviations, at high densities, from the law of Guldberg and Waage as to binary collisions is therefore, from (19) and (12),

$$\frac{1}{1-(n_1+n_2)\sigma^3 2^{-\frac{1}{2}}} \exp\left\{\frac{-0{\cdot}79(n_1+n_2)\omega}{1-(n_1+n_2)\sigma^3 2^{-\frac{1}{2}}}\right\} = F(n_1+n_2), \quad \text{say.} \tag{21}$$

The numerical values of $F(n_1+n_2)$ are shown below as a function of the ratio of the total density to the density at N.T.P. The numerical data are $\sigma = 3\times 10^{-8}$ cm., $\omega = 6{\cdot}65\times 10^{-23}$ cm.^{3}, as before; together with $n_1+n_2 = 2{\cdot}7\times 10^{19}$ cm.$^{-3}$ at N.T.P.

density as ratio to that at N.T.P.	10^{-1}	1	10	10^2	10^3
$F(n_1+n_2)$	0·9999	0·9991	0·991	0·908	0·110

At low densities one may expect agreement with the classical theory. But, because the measure of encounters is different, the comparison is somewhat indirect. Jeans (1916, p. 37) gives for the number of binary collisions per cm.^{3} and per second

$$\pi n^2 \sigma^2 \bar{c} 2^{-\frac{1}{2}}, \tag{22}$$

where $\bar{c}$ is the mean velocity, equal to 0·921 of the root-mean-square velocity C. The average duration of a single encounter equals the ratio of the aggregate duration of all encounters to the number of such encounters. So from (22) and from (19) at low densities,

$$\{\text{Average duration of an encounter}\} = \frac{\frac{1}{2}n^2\omega}{\pi n^2\sigma^2\bar{c}2^{-\frac{1}{2}}} = \frac{1{\cdot}661\times 10^{-8}}{\bar{c}} \text{ sec.} \tag{23}$$

because, by (9), $\omega = 2{\cdot}46\sigma^3$ cm.^{3}.

The purpose of Jeans in the passage already quoted was to estimate the duration of a collision; and he continued with the following words:

> 'If the molecules had initially a relative velocity V, this acute part of the collision may be supposed described with an average relative velocity $\frac{1}{2}V$, and so we obtain for the time of the half-collision about $10^{-8}/V$ seconds.' (24)

The expressions (23) and (24) are of the same order of magnitude, so that rough agreement is verified. But it is noticeable that the connexion has to be made to one of the seldom mentioned and least definite results of the classical theory.

It is also noteworthy that the aggregate duration of all binary encounters does not depend on the temperature nor on the molecular mass, whereas the average duration of a single binary encounter varies with both of them.

At high densities the effect of the factor $\Omega\exp\{-0{\cdot}79(n_1+n_2)\Omega\}$ in (19) will presumably be to make the rate of bimolecular chemical reaction fall below that which would be expected from the law of Guldberg and Waage. Effects more or less of that sort have been observed, and have puzzled observers; for Hinshelwood (1940, p. 28) mentioned '...the uncertainty whether a second-order reaction is really bimolecular, or whether it may prove at higher pressures to become of the first order and so reveal itself as essentially unimolecular....' Possibly the present theory may explain some of those effects as quite proper to genuinely bimolecular reactions. But the phenomena may be complicated by activation, by the walls of the vessel, by impurities in the gas, by delayed actions, by intermediate products, and by chain reactions, so that I had better leave any application of the present theory to be made by those who are in close contact with the experiments. There is also the complication caused by ternary collisions at high densities. For the error in the customary theories is their inclusion, in the binary class, of encounters that involve more than two molecules. Yet these may, perhaps, have the same chemical effect as binary encounters.

Ternary encounters

A simple theory by Bodenstein (Hinshelwood 1940, p. 142) has already been mentioned. Elaborate investigations have been published by W. Steiner (1932) and by R. H. Fowler (1936, pp. 691–5). The last two are in many respects far more thorough than the present note. Yet none of those three publications alludes to the probability that the extraneous molecules will stay away. So there is something more to be said.

The number of ternary encounters existing simultaneously in a cm.^{3} was estimated by the first theory to be of the order of

$$\tfrac{1}{6}n^3\omega_1^2 e^{-n\omega_1}. \tag{25}$$

Now make a revised estimate. There are several different varieties of ternary encounter. The expression (25) includes them all. But for chemical purposes it may be more suitable to select those in which a molecule of the first species encounters two molecules of the second species, and to ignore all the other ternary encounters. Even in the selected group there are two varieties. For definiteness select

> The variety in which a molecule of one species is in encounter with two molecules of a different species, and these like molecules are in encounter with each other. (26)

The centre B_2 of the third molecule is thus restricted to lie in a ring cut from the spherical shell. In the diagram, A_2 and B_2 are shown in their mean positions, and the cross-section of the ring is shaded. The volume of the ring was found graphically by Pappus's theorem to be

$$0{\cdot}067\omega. \tag{27}$$

The probability that A_2 lies in the shell is Ω as before. When A_2 is in the shell, the probability that B_2 is in the ring is $0{\cdot}067\Omega$. That completes the required bunch of three molecules. But we have also to consider the probability that an extraneous molecule may have its centre, C say, in the vacant part of the shell. In the diagram the cones of exclusion connected with A_2 and B_2 appear in section as radii; that for A_2 is drawn with full lines, that for B_2 with broken lines. The cones overlap; the angle between their axes being, in the mean position, $50° + 5° + 5° = 60°$. The fraction of the spherical shell, left available for any other molecular centre C, is evidently, from (13), more than $1 - 2 \times 0{\cdot}213$ and less than $1 - 0{\cdot}213$. After some spherical trigonometry the correct fraction was found to be 0·652, in the mean positions of A_2 and B_2. The rest of the argument is similar to that already set out for binary encounters. The result is that, in a cm.³ and during 1 sec. of the observer's time, the aggregate duration of ternary encounters belonging to the variety (26) is

$$\tfrac{1}{2} n_1 n_2^2\, 0{\cdot}067 \Omega^2 \exp\{-0{\cdot}65(n_1+n_2)\Omega\}. \tag{28}$$

For comparison with (25) it is useful to note that $\frac{1}{2} n_1 n_2^2$ is a term in the expansion of $\frac{1}{6}(n_1+n_2)^3$.

For another variety of ternary encounter in which the two like molecules do not encounter each other, the calculation of the excluded volumes would be more complicated.

Similar arguments could presumably be extended to four or five molecules in a bunch, but with increasing complication.

CRITICISMS AND POSSIBLE IMPROVEMENTS

The main objection to both the foregoing theories is that their standard volumes, ω_1 and ω, are not well defined. They are known within a factor of three, but not to 10 %. Indeed, ω_1 has already been abandoned in favour of ω. That objection should not be directed against the foregoing second theory, nor against Jeans's estimate of the duration of a collision, but against the concept of encounter itself; for the concept is somewhat vague. It attains definiteness only when the encounters are imagined to be instantaneous collisions between hard bodies; but no one believes that molecules are really like that.

A possible improvement would begin by subdividing the concept of encounter. Instead of saying that two molecules were in encounter, one could say that the distance apart of their centres lay between r and $r+dr$. The vague volume ω would thus be replaced by $4\pi r^2 dr$. The foregoing arguments about probability would apply to the thin spherical shell. There would be an opportunity to replace the rough exclusion factor 0·79 by the correct function of r, and to allow for the effect of molecular attraction in increasing the average number of molecules in the shell, after the manner of Boltzmann (Jeans 1916, p. 83). When all that had been written in detail, it might be so elaborate that chemists would not read it.

Another, and probably more generally acceptable, way of escape from the vagueness of ω, is to regard ω as an empirical constant which may perhaps be determined by experiment for particular substances.

([*Note added* 2 *August* 1945.] Theories I and II both assume that the distribution of the molecules is statistically steady. Any chemical process is a departure from steadiness. So we must expect that Theories I and II will describe slow reactions, but not explosions.

The present paper, taken as a whole, calls for some revision of the classical theory of collisions. The classical accurate theory necessitates finding the velocity distribution; so that, to lighten the burden, unnatural simplifications have sometimes been made elsewhere. For example Chapman and Cowling, in their standard treatise (1939), devoted a chapter to Enskog's theory of dense gases; but for simplicity Enskog assumed (p. 273) collisions to be instantaneous, so that multiple encounters were negligible. In so doing he ignored the main problem to which the present paper offers a solution. It is remarkable that Theories I and II arrive at results without any need to specify the velocity distribution. The absence of that classical complication sets the mind free to consider other probabilities.)

REFERENCES

Aitken, A. C. 1942 *Statistical mathematics,* 2nd ed. Edinburgh: Oliver and Boyd.
Bragg, W. L. 1933 *The crystalline state,* **1**. London: Bell and Sons.
Bridgman, P. W. 1931 *The physics of high pressure.* London: Bell and Sons.
Chapman, S. & Cowling, T. G. 1939 *The mathematical theory of non-uniform gases.* Cambridge: University Press.
Fowler, R. H. 1936 *Statistical mechanics,* 2nd ed. Cambridge: University Press.
Gosset, W. S. (alias 'Student') 1907 *Biometrika,* **5**, 351.
Hinshelwood, C. N. 1940 *The kinetics of chemical change.* Oxford: Clarendon Press.
Jeans, J. H. 1916 *The dynamical theory of gases,* 2nd ed. Cambridge: University Press.
Sheppard, W. F. 1929 *Encyc. Brit.* 14th ed. **18**, 536.
Steiner, W. 1932 *Z. phys. Chem.* **15**, 249–273.

On the conservation of momentum

By E. A. Milne, F.R.S.

(*Received* 22 *August* 1945)

1. *Scope of the paper.* The present paper is a continuation of an earlier sequence in these *Proceedings*, entitled 'The inverse square law of gravitation, I, II and III' (Milne 1936, 1937*c*), which itself was connected with another sequence entitled 'Kinematics, dynamics and the scale of time, I, II and III' (Milne 1937*a*, *b*). In the latter sequence the relations between t-dynamics and τ-dynamics were obtained. In the former sequence a formula for the gravitational potential energy of two particles was isolated which was Lorentz-invariant for transformations from any one fundamental observer in the expanding substratum to any other fundamental observer; and the corresponding equations of motion in t-measure were obtained. In the present paper I obtain relations which correspond to the integrals of linear and angular momentum in the many-body problem of classical gravitational theory. I conclude that Einstein's form of the conservation of linear momentum in special relativity holds good in the universe at large only in the case of a collinear set of particles moving along their line of collinearity, and I give the modification of the law of conservation which should hold good in other cases.

2. *Nature of the solution.* A fundamental step in the solution of the many-body problem in classical gravitational theory is the isolation of the integrals of linear and angular momentum. Either these can be written down at sight, on the basis of an assumed principle of conservation of momentum, or they can be deduced from the equations of motion. In kinematic relativity, where laws of nature are not assumed and where they take unfamiliar forms, it would be contrary to the spirit of the subject to *assume* forms for the principles of the conservation of momentum, and one is driven to attempt to deduce the integrals in question from the kinematically derived equations of motion. In the classical case, the essence of the step involved in obtaining these integrals is the use of an identity satisfied by the gravitational potential, by means of which an integrable combination of the equations of motion is obtained from which the gravitational potential is absent. In one of my earlier papers already cited (Gravitation, II, p. 12) I gave an identity satisfied by χ, the gravitational potential in kinematic relativity which I stated should be capable of yielding a relation corresponding to the usual integral of *angular* momentum; but no identity was given which would correspond to the integral of *linear* momentum. The identity requisite for this purpose I give in this paper.

3. *Difficulties in the concept of momentum.* Before proceeding, some difficulties in the concept of linear momentum may be noticed.* In general relativity, it is usually stated that the vanishing of the divergence of

$$G^{\nu}_{\mu}-\tfrac{1}{2}g^{\nu}_{\mu}G$$

---

* Similar difficulties occur in the concept of angular momentum, but for simplicity the arguments given here refer explicitly to linear momentum.

is equivalent to the conservation laws of dynamics (Eddington 1923, p. 115). The identities satisfied do in fact correspond to the conservation laws of dynamics in the sense in which they are applied to the transport of energy and momentum in a *fluid*. But they are not helpful in suggesting the form taken by the conservation laws in the case of a set of discrete *particles*; and, in fact, little progress has been made, in general relativity, with the solution of the problem of even two gravitating particles (Eddington 1923, p. 95).

In special relativity, Einstein's expression $m\mathbf{v}/(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}$ for the linear momentum of a particle of rest-mass m, moving with velocity v, leads us to expect that for any system of interacting particles, the vector sum

$$\frac{m_1\mathbf{v}_1}{(1-\mathbf{v}_1^2/c^2)^{\frac{1}{2}}}+\frac{m_2\mathbf{v}_2}{(1-\mathbf{v}_2^2/c^2)^{\frac{1}{2}}}+\dots \tag{1}$$

will be conserved. This is unambiguous when, for example, the only interactions take the form of collisions, between which the velocities of the separate particles remain *constant*; for then, provided one remains inside an interval bounded by two consecutive collisions, it is immaterial at what instant one evaluates the velocity of each particle. But if the interaction is continuous, as in the case of gravitating or electrically interacting particles, the velocities to be assigned to the various particles in evaluating the total momentum depend on the epochs at which these velocities are evaluated; for the velocity of each particle is continuously altering in time. Some convention as to simultaneity is accordingly required before a statement of the conservation of momentum can have a meaning: it is necessary to state a convention fixing the instants at which the separate terms of (1) are to be evaluated. But events which are simultaneous for one observer, on any particular convention, will not necessarily be simultaneous for another observer, on the same convention. Accordingly (1) cannot be expected to be conserved, in the many-body gravitational problem, without further precision of the meaning of the evaluation of each term. I shall show that in kinematic relativity, when the τ-measure of time is employed, expression (1) is conserved for all fundamental observers (who have, of course, in τ-measure a world-wide simultaneity) provided it relates to a *collinear* set of particles, i.e. to a set for which the accelerations are parallel to the velocities; but when the accelerations are not parallel to the velocities, expression (1) is not conserved; and I shall obtain its rate of change. I shall, however, in the first place conduct the investigation in t-measure, as t-measure is the more fundamental.

4. *Equations of motion in t-measure.* The equations of motion, in t-measure, of a gravitating set of particles of masses $m_1, m_2, \dots, m_n$, are derived in kinematic relativity from a gravitational potential χ, defined by

$$\chi = \sum_{\mu\neq\nu} \chi_{\mu\nu} \quad (\mu, \nu = 1, 2, \dots, n),$$

where, for example,
$$\chi_{12} = -\frac{m_1 m_2 c^2}{M}\frac{X_{12}}{(X_{12}^2 - X_1 X_2)^{\frac{1}{2}}}, \tag{2}$$

and
$$X_1 = t_1^2 - \mathbf{P}_1^2/c^2, \quad X_2 = t_2^2 - \mathbf{P}_2^2/c^2, \quad X_{12} = t_1 t_2 - \mathbf{P}_1 . \mathbf{P}_2/c^2,$$

$\mathbf{P}_1, \mathbf{P}_2, \ldots$ being the position vectors of the particles $m_1, m_2, \ldots$ with reference to any fundamental observer O (a member of the substratum) at the epochs $t_1, t_2, \ldots$ reckoned from the natural zero of time, and M being the mass of the equivalent homogeneous universe (Gravitation, II, p. 7). With this definition of χ, the equations of motion of the particle m_1 are

$$\frac{1}{Y_1^{\frac12}}\frac{d}{dt_1}\left[m_1\xi_1^{\frac12}\frac{\mathbf{V}_1}{Y_1^{\frac12}}\right] = -\frac{m_1\xi_1^{\frac12}}{X_1}\left(\mathbf{P}_1-\mathbf{V}_1\frac{Z_1}{Y_1}\right)+2\frac{\mathbf{V}_1}{Y_1^{\frac12}}\frac{1}{Y_1^{\frac12}}\frac{d}{dt_1}(m_1\xi_1^{\frac12})-\frac{\partial\chi}{\partial\mathbf{P}_1}, \tag{3}$$

$$\frac{1}{Y_1^{\frac12}}\frac{d}{dt_1}\left[m_1\xi_1^{\frac12}\frac{c}{Y_1^{\frac12}}\right] = -\frac{m_1\xi_1^{\frac12}}{X_1}\left(ct_1-c\frac{Z_1}{Y_1}\right)+2\frac{c}{Y_1^{\frac12}}\frac{1}{Y_1^{\frac12}}\frac{d}{dt_1}(m_1\xi_1^{\frac12})+\frac{1}{c}\frac{\partial\chi}{\partial t_1}. \tag{3'}$$

Here $\qquad Y_1 = 1-\mathbf{V}_1^2/c^2, \quad Z_1 = t_1-\mathbf{P}_1.\mathbf{V}_1/c^2, \quad \xi_1 = Z_1^2/X_1Y_1.$

$\mathbf{V}_1$ is the velocity of the particle m_1 relative to the observer O. Equation (3′) is a consequence of (3) in virtue of the identity

$$t_1\frac{\partial\chi}{\partial t_1}+\mathbf{P}_1.\frac{\partial\chi}{\partial\mathbf{P}_1} = 0.$$

Equation (3′) is, of course, the time-component corresponding to the three-vector equation (3). In (3) and (3′), $m_1\xi_1^{\frac12}$ represents the mass of the moving particle m_1, and the left-hand side of (3) represents the rate of change of its momentum relative to the arbitrary observer O; the first term on the right-hand side represents the pull of the substratum, in t-measure; the second term on the right represents the effect of change of mass with velocity.

For each particle of the system there is a similar pair of equations of motion. This system of equations does not become definite until relations are imposed connecting $t_1, t_2, \ldots, t_n$ with one another. I originally stated that the set of equations was to be solved with the simultaneity convention

$$t_1 = t_2 = \ldots = t_n = t,$$

but it was pointed out by Schild (1944) that this would not yield Lorentz-invariant orbits. In my applications, as recognized by Schild, no errors were actually committed, for in the application to two particles I took the origin O at one of the particles, m_2, and since then χ_{12} depends explicitly only on t_1, t_1 can be used as an independent variable without needing any simultaneity convention. It has recently been pointed out by Camm (1945) that the only possible simultaneity convention yielding Lorentz-invariant orbits is

$$X_1 = X_2 = \ldots = X_n. \tag{4}$$

This has the further merit, as pointed out by Whitrow (1945), that it corresponds in τ-measure to the convention

$$\tau_1 = \tau_2 = \ldots = \tau_n = \tau, \tag{4'}$$

which is consistent with the absolute simultaneity which is characteristic of the substratum when described in τ-measure. I therefore adopt Camm's convention (4) in connexion with the set of equations (3), (3′).

But the *relations* between the solutions of the set (3), (3′) which correspond to the classical integrals of linear and angular momentum will be found to be independent of any simultaneity convention. A simultaneity convention is needed only when one integrates these relations. It was pointed out by Camm that a similar independence of simultaneity convention holds with respect to the energy integral of the system, though in this case the independence holds even for the operation of integration. In all cases, of course, a simultaneity convention is needed when we come to interpret the integrals numerically.

5. *The fundamental identity corresponding to the integral of linear momentum.* One looks now for a vector combination of the set of equations of motion of the type (3), (3′) which is independent of χ. For this purpose a vector identity satisfied by each $\chi_{\mu\nu}$ is needed.

Consider the space and time derivatives of χ_{12}. By direct differentiation

$$-\frac{\partial\chi_{12}}{\partial\mathbf{P}_1}=\frac{m_1m_2}{M}\frac{X_2(X_1\mathbf{P}_2-X_{12}\mathbf{P}_1)}{(X_{12}^2-X_1X_2)^{\frac{3}{2}}},\quad \frac{\partial\chi_{12}}{c\,\partial t_1}=\frac{m_1m_2c}{M}\frac{X_2(X_1t_2-X_{12}t_1)}{(X_{12}^2-X_1X_2)^{\frac{3}{2}}}.$$

Combination of these gives

$$-t_1\frac{\partial\chi_{12}}{\partial\mathbf{P}_1}-\frac{\mathbf{P}_1}{c^2}\frac{\partial\chi_{12}}{\partial t_1}=\frac{m_1m_2}{M}\frac{X_2X_1}{(X_{12}^2-X_1X_2)^{\frac{3}{2}}}(t_1\mathbf{P}_2-t_2\mathbf{P}_1).$$

Interchanging the suffixes 1 and 2, we have

$$-t_2\frac{\partial\chi_{12}}{\partial\mathbf{P}_2}-\frac{\mathbf{P}_2}{c^2}\frac{\partial\chi_{12}}{\partial t_2}=\frac{m_1m_2}{M}\frac{X_2X_1}{(X_{12}^2-X_1X_2)^{\frac{3}{2}}}(t_2\mathbf{P}_1-t_1\mathbf{P}_2).$$

The right-hand sides being antisymmetrical in the suffixes 1 and 2, we have, on addition,

$$-t_1\frac{\partial\chi_{12}}{\partial\mathbf{P}_1}-\frac{\mathbf{P}_1}{c^2}\frac{\partial\chi_{12}}{\partial t_1}-t_2\frac{\partial\chi_{12}}{\partial\mathbf{P}_2}-\frac{\mathbf{P}_2}{c^2}\frac{\partial\chi_{12}}{\partial t_2}=0. \tag{5}$$

A similar identity is satisfied by each $\chi_{\mu\nu}$. It will be found that these are the desired identities.

6. *The relation corresponding to the principle of linear momentum.* Multiplying (3) by t_1 and (3′) by $-\mathbf{P}_1/c$, and adding together similar multiples of each pair of the set, it is seen that in virtue of identities (5) the gravitational potential is eliminated, and one is left with the three-vector relation

$$\Sigma\left[\frac{m_1\xi_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}}\left[t_1\frac{d}{dt_1}\left(\frac{\mathbf{V}_1}{Y_1^{\frac{1}{2}}}\right)-\mathbf{P}_1\frac{d}{dt_1}\left(\frac{1}{Y_1^{\frac{1}{2}}}\right)\right]\right.$$
$$\left.+\frac{m_1\xi_1^{\frac{1}{2}}}{X_1}\frac{Z_1}{Y_1}(\mathbf{P}_1-\mathbf{V}_1t_1)-\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Y_1^{\frac{1}{2}}}\frac{1}{Y_1^{\frac{1}{2}}}\frac{d}{dt_1}(m_1\xi_1^{\frac{1}{2}})\right]=0, \tag{6}$$

where the summation is extended to all suffixes $1, 2, \ldots, n$.

Each member of the summation is readily simplified. We have, since $d\mathbf{P}_1/dt_1 = \mathbf{V}_1$,

$$t_1\frac{d}{dt_1}\left(\frac{\mathbf{V}_1}{Y_1^{\frac{1}{2}}}\right)-\mathbf{P}_1\frac{d}{dt_1}\left(\frac{1}{Y_1^{\frac{1}{2}}}\right)=\frac{d}{dt_1}\left(\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Y_1^{\frac{1}{2}}}\right),$$

and since $dX_1/dt_1 = 2Z_1$, we have

$$-\frac{m_1\xi_1^{\frac{1}{2}}}{X_1}\frac{Z_1}{Y_1}(\mathbf{V}_1t_1-\mathbf{P}_1)=-m_1\xi_1^{\frac{1}{2}}\left(\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Y_1^{\frac{1}{2}}}\right)\frac{1}{Y_1^{\frac{1}{2}}}\frac{d}{dt_1}(\log X_1^{\frac{1}{2}}).$$

Hence the typical term in the summation (6) is

$$\frac{m_1\xi_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}}\left[\frac{d}{dt_1}\left(\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Y_1^{\frac{1}{2}}}\right)-\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Y_1^{\frac{1}{2}}}\frac{d}{dt_1}(\log X_1^{\frac{1}{2}}+\log m_1\xi_1^{\frac{1}{2}})\right].$$

The integrating factor of the [...] is

$$\exp\left[-\int^{t_1}\frac{d}{dt_1}(\log X_1^{\frac{1}{2}}\xi_1^{\frac{1}{2}})\,dt_1\right],$$

i.e.

$$X_1^{-\frac{1}{2}}\xi_1^{-\frac{1}{2}}.$$

Hence the summation (6) may be written

$$\Sigma X_1^{\frac{1}{2}}\xi_1^{\frac{1}{2}}\frac{m_1\xi_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}}\frac{d}{dt_1}\left[\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Y_1^{\frac{1}{2}}X_1^{\frac{1}{2}}\xi_1^{\frac{1}{2}}}\right]=0,$$

or

$$\Sigma m_1\frac{\xi_1X_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}}\frac{d}{dt_1}\left[\frac{\mathbf{V}_1t_1-\mathbf{P}_1}{Z_1}\right]=0. \tag{7}$$

7. *Physical recognition.* Relation (7) is readily recognized as the differential form of the integral of linear momentum in t-measure. It will be transformed rigorously to τ-measure in a moment, but to see its physical meaning in a simple way we write it down restricting attention to distances $|\mathbf{P}|$ small compared with ct, the radius of the universe, and to velocities $|\mathbf{V}|$ small compared with c. Then, approximately,

$$\Sigma m_1t_1\frac{d}{dt_1}\left(\mathbf{V}_1-\frac{\mathbf{P}_1}{t_1}\right)=0.$$

But $\mathbf{V}_1-\mathbf{P}_1/t_1$ is the velocity of m_1 relative to the fundamental observer past whom m_1 is instantaneously passing; for $\mathbf{P}_1/t_1$ is the expansion velocity of the substratum at m_1. Hence $\mathbf{V}_1-\mathbf{P}_1/t_1$ is the absolute velocity $\mathbf{v}_1$, in τ-measure, relative to the substratum, which of course in τ-measure is relatively stationary. Also since the clock regraduation formula from t-time to τ-time is

$$\frac{dt}{t}=\frac{d\tau}{t_0},$$

t_1d/dt_1 is just $t_0d/d\tau_1$. Thus the approximate form of (7) in τ-measure is

$$\Sigma m_1\frac{d\mathbf{v}_1}{d\tau_1}=0.$$

This relation has been derived without imposing a specification of a simultaneity convention. If now (4) be imposed in its τ-form (4′), then

$$\Sigma \frac{m_1 d\mathbf{v}_1}{d\tau} = 0,$$

which integrates to $$\Sigma m_1 \mathbf{v}_1 = \text{const.} \tag{8}$$

This is sufficient to identify the exact relation (7) as the counterpart in t-dynamics of the integral of linear momentum in classical dynamics.

8. *Angular momentum.* Before going further we obtain the similar relation which corresponds to the integral of angular momentum. The necessary identity, easily verified, is

$$\mathbf{P}_1 \wedge \left(-\frac{\partial \chi_{12}}{\partial \mathbf{P}_1}\right) + \mathbf{P}_2 \wedge \left(-\frac{\partial \chi_{12}}{\partial \mathbf{P}_2}\right) = 0. \tag{9}$$

Multiplying (3) vectorially by $\mathbf{P}_1$, and adding n similar equations, we find

$$\Sigma \left[\frac{m_1 \xi_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}} \frac{d}{dt_1}\left(\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}}\right) - \frac{m_1 \xi_1^{\frac{1}{2}}}{X_1} \frac{Z_1}{Y_1} (\mathbf{P}_1 \wedge \mathbf{V}_1) - \frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}} \frac{1}{Y_1^{\frac{1}{2}}} \frac{d}{dt_1} (m_1 \xi_1^{\frac{1}{2}})\right] = 0,$$

the summation as before being extended to suffixes $1, 2, \ldots, n$. This may be written

$$\Sigma \frac{m_1 \xi_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}} \left[\frac{d}{dt_1}\left(\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}}\right) - \frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}} \left(\frac{Z_1}{X_1} + \frac{d}{dt_1} \log m_1 \xi_1^{\frac{1}{2}}\right)\right] = 0.$$

As before, $$\frac{Z_1}{X_1} = \frac{d}{dt_1} \log X_1^{\frac{1}{2}},$$

and the 'integrating factor' of the [...] is again $X_1^{-\frac{1}{2}} \xi_1^{-\frac{1}{2}}$. Hence

$$\Sigma X_1^{\frac{1}{2}} \xi_1^{\frac{1}{2}} \frac{m_1 \xi_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}} \frac{d}{dt_1} \left[\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}} X_1^{\frac{1}{2}} \xi_1^{\frac{1}{2}}}\right] = 0,$$

or $$\Sigma m_1 \frac{\xi_1 X_1^{\frac{1}{2}}}{Y_1^{\frac{1}{2}}} \frac{d}{dt_1} \left[\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Z_1}\right] = 0. \tag{10}$$

The approximate form of (10) in τ-measure is

$$\Sigma m_1 \frac{d}{d\tau_1} (\mathbf{\Pi}_1 \wedge \mathbf{v}_1) = 0,$$

where $\Pi_1 = t_0 P_1 / t_1$ is the τ-measure of the radius vector. With simultaneity convention (4′), this integrates to

$$\Sigma m_1 (\mathbf{\Pi}_1 \wedge \mathbf{v}_1) = \text{const.}, \tag{11}$$

which is sufficient to identify (10) as the relation in t-dynamics corresponding to the integral of angular momentum in classical dynamics. For a single particle in the vicinity of a massive nucleus (Keplerian problem), (10) gives

$$m_1 (\mathbf{P}_1 \wedge \mathbf{V}_1) \propto Z_1 \sim t_1,$$

which is the usual secular increase of angular momentum found previously (Gravitation, III, p. 32).

9. It is to be noted that both (7) and (10) consist of the sum of a number of functions of the separate epochs, positions and velocities of the separate particles $m_1, m_2, \ldots, m_n$, independent of whatever simultaneity convention connecting $t_1, t_2, \ldots, t_n$ is imposed in solving the equations of motion. The actual numerical values to be assigned to the separate terms of each summation, during the motion, depend of course on the simultaneity convention adopted, but the *form* of the relationship is independent of the convention. The existence of relationships of this kind is a deep-lying property of the t-dynamics, and it may be considered as the ultimate reason for the recognition of conservation laws in the universe at large. It would be of great interest to investigate whether similar relations can be deduced for a number of interacting *charged* particles, using the rational electrodynamics developed previously (Milne 1938, 1943).

10. *Momentum as a six-vector.* In passing it will be noted that

$$\left(\frac{\mathbf{V}_1 t_1 - \mathbf{P}_1}{Y_1^{\frac{1}{2}}},\quad \frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}}\right)$$

is a six-vector, or antisymmetrical tensor of rank 2, formed out of the four-vectors $(\mathbf{P}_1, ct_1)$ and $(\mathbf{V}_1/Y_1^{\frac{1}{2}}, c/Y_1^{\frac{1}{2}})$. Since $Z_1/Y_1^{\frac{1}{2}}$ is an invariant,

$$\left(\frac{\mathbf{V}_1 t_1 - \mathbf{P}_1}{Z_1},\quad \frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Z_1}\right)$$

is also a six-vector. It follows that the linear and angular momentum of a set of particles, on a suitable simultaneity convention, together constitute a six-vector in t-measure.

11. *Transformation to τ-measure.* Now transform (7) and (10) to τ-measure, without the limitation to small velocities, but retaining the limitation to distances small compared with the radius of the universe. The reason for retaining this second limitation is that then the hyperbolic space of τ-measure may be considered as sensibly Euclidean.

When all observers' clocks are regraduated from t to τ, where

$$\tau = t_0 \log (t/t_0) + t_0,$$

the corresponding transformation of co-ordinates is

$$t = t_0 e^{(\tau - t_0)/t_0} \cosh |\mathbf{\Pi}|/ct_0, \quad |\mathbf{P}|/c = t_0 e^{(\tau - t_0)/t_0} \sinh |\mathbf{\Pi}|/ct_0$$

(Scale of time, II, p. 177). Also the metric $d\sigma$ of the public hyperbolic space used instead of the t-observers' private Euclidean spaces is given by

$$Y^{\frac{1}{2}} dt = ds = e^{(\tau - t_0)/t_0} d\sigma.$$

The transformation gives also $\quad X^{\frac{1}{2}} = t_0 e^{(\tau - t_0)/t_0},$

and $$\xi^{\frac{1}{2}} = \frac{Z}{X^{\frac{1}{2}} Y^{\frac{1}{2}}} = \frac{1}{Y^{\frac{1}{2}}} \frac{dX^{\frac{1}{2}}}{dt} = \frac{t_0 d(e^{(\tau - t_0)/t_0})}{e^{(\tau - t_0)/t_0} d\sigma} = \frac{d\tau}{d\sigma} = \left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{-\frac{1}{2}}.$$

For $|\mathbf{\Pi}| \ll ct_0$, we have $$\mathbf{P} = (t/t_0)\mathbf{\Pi},$$

whence $$\mathbf{V} = \frac{d\mathbf{P}}{dt} = \frac{\mathbf{\Pi}}{t_0} + \frac{t}{t_0}\frac{d\mathbf{\Pi}}{dt} = \frac{\mathbf{P}}{t} + \frac{d\mathbf{\Pi}}{d\tau},$$

and so $$\frac{\mathbf{V}t - \mathbf{P}}{Z} \sim \frac{\mathbf{V}t - \mathbf{P}}{t} = \frac{d\mathbf{\Pi}}{d\tau} = \mathbf{v}.$$

These relations are sufficient to transform (7), and hence

$$\Sigma m_1 \left(1 - \frac{\mathbf{v}_1^2}{c^2}\right)^{-1} \frac{d\mathbf{v}_1}{d\sigma} = 0,$$

or $$\Sigma m_1 \left(1 - \frac{\mathbf{v}_1^2}{c^2}\right)^{-\frac{3}{2}} \frac{d\mathbf{v}_1}{d\tau_1} = 0. \tag{12}$$

This is the relation in τ-measure corresponding to the integral of linear momentum in classical mechanics. It should hold for unrestricted velocities. Similarly, the τ-relation corresponding to the integral of angular momentum in classical mechanics is found to be, from (10),

$$\Sigma m_1 \left(1 - \frac{\mathbf{v}_1^2}{c^2}\right)^{-\frac{3}{2}} \frac{d}{d\tau_1}(\mathbf{\Pi}_1 \wedge \mathbf{v}_1) = 0. \tag{13}$$

These are to be interpreted with the simultaneity convention (4′).

12. *Comparison with special relativity.* It is well known that if v is a *scalar* variable, then

$$\frac{d}{d\tau}\left[\frac{v}{(1 - v^2/c^2)^{\frac{1}{2}}}\right] = \frac{dv/d\tau}{(1 - v^2/c^2)^{\frac{3}{2}}}. \tag{14}$$

Consequently for a set of collinear particles in motion along their line of collinearity, (12) integrates at once to the form

$$\Sigma \frac{m_1 v_1}{(1 - v_1^2/c^2)^{\frac{1}{2}}} = \text{const.} \tag{15}$$

This is Einstein's special relativity form of the conservation of linear momentum, already discussed in § 3.

But when $\mathbf{v}$ is a vector variable, then

$$\begin{aligned}\frac{d}{d\tau}\left[\frac{\mathbf{v}}{(1 - \mathbf{v}^2/c^2)^{\frac{1}{2}}}\right] &= \frac{d\mathbf{v}/d\tau}{(1 - \mathbf{v}^2/c^2)^{\frac{1}{2}}} + \frac{\mathbf{v}}{c^2}\frac{\mathbf{v}.d\mathbf{v}/d\tau}{(1 - \mathbf{v}^2/c^2)^{\frac{3}{2}}} \\ &= \frac{d\mathbf{v}/d\tau}{(1 - \mathbf{v}^2/c^2)^{\frac{3}{2}}} + \frac{1}{c^2}\frac{-\mathbf{v}^2 d\mathbf{v}/d\tau + (\mathbf{v}.d\mathbf{v}/d\tau)\mathbf{v}}{(1 - v^2/c^2)^{\frac{3}{2}}} \\ &= \frac{d\mathbf{v}/d\tau}{(1 - \mathbf{v}^2/c^2)^{\frac{3}{2}}} + \frac{1}{c^2}\frac{(d\mathbf{v}/d\tau \wedge \mathbf{v}) \wedge \mathbf{v}}{(1 - \mathbf{v}^2/c^2)^{\frac{3}{2}}}.\end{aligned}$$

Hence, *in general*, if the vector linear momentum $\mathbf{L}$ be defined by the left-hand side of (1), relation (12) yields

$$\frac{d\mathbf{L}}{d\tau} = \frac{1}{c^2}\Sigma m_1 \frac{(d\mathbf{v}_1/d\tau \wedge \mathbf{v}_1) \wedge \mathbf{v}_1}{(1 - v_1^2/c^2)^{\frac{3}{2}}}. \tag{16}$$

This shows that L is constant only when each acceleration $d\mathbf{v}/d\tau$ is parallel to the corresponding velocity $\mathbf{v}_1$, which, as before stated, will only occur for a collinear set of particles in motion along their line of collinearity.

Similarly, if one defines the angular momentum $\mathbf{H}$ by

$$\mathbf{H} = \Sigma m_1 \frac{\mathbf{\Pi}_1 \wedge \mathbf{v}_1}{(1-\mathbf{v}_1^2/c^2)^{\frac{1}{2}}}, \tag{17}$$

then (13) gives

$$\frac{d\mathbf{H}}{d\tau} = \frac{1}{c^2} \Sigma m_1 \mathbf{\Pi}_1 \wedge \frac{(d\mathbf{v}_1/d\tau \wedge \mathbf{v}_1) \wedge \mathbf{v}_1}{(1-\mathbf{v}_1^2)/c^2)^{\frac{3}{2}}}. \tag{18}$$

It may be concluded, on the basis of the present formulation of dynamics, that linear and angular momentum for a gravitationally interacting set of particles, as these momenta are defined in special relativity, are not in general conserved. Further, since the right-hand sides of (16) and (18) are not complete differential coefficients, it is not possible to redefine linear or angular momentum in such a way as to make them conserved.

13. *Alternative derivation.* An alternative procedure is to begin with the approximate form of the gravitational equations of motion in τ-measure, and obtain the relations corresponding to the classical integrals directly. It has been shown (Gravitation, III, p. 29) that in τ-measure the approximate gravitational equations are of the form

$$\frac{m_1 d\mathbf{v}_1/\partial\tau_1}{(1-\mathbf{v}_1^2/c^2)^{\frac{3}{2}}} = -\frac{\partial\chi}{\partial\mathbf{\Pi}_1}, \tag{19}$$

χ being expressed in τ-measure. Since χ_{12} is a function of $|\mathbf{\Pi}_1 - \mathbf{\Pi}_2|$, we have

$$\frac{\partial\chi_{12}}{\partial\mathbf{\Pi}_1} = -\frac{\partial\chi_{12}}{\partial\mathbf{\Pi}_2}, \tag{20}$$

and hence addition of the complete set of equations as they stand eliminates χ and yields just (12). This method of proceeding fails, however, to yield the exact relation (7) holding in t-measure.

14. It may perhaps be objected that (19) must be wrong, because it is not of the form that would be expected from Einstein's special relativity dynamics: it might have been expected that

$$m_1 \frac{d}{d\tau_1} \frac{\mathbf{v}_1}{(1-\mathbf{v}_1^2/c^2)^{\frac{1}{2}}} = -\frac{\partial\chi}{\partial\mathbf{\Pi}_1}. \tag{21}$$

But there is no justification for this expectation, other than the supposition that the left-hand side *ought* to be of the form of the derivative of the momentum as defined by Einstein. The form (19), on the other hand, emerges as the result of a long train of argument, beginning in t-measure, and not depending on any preconceived form of the equations of motion. Most certainly (19) was not derived by any teleological train of argument having Einstein's special relativity dynamics, or something akin to it, as an ultimate destination. It may be pointed out that both (19) and (21), each taken with the other members of its set, possess the energy integral

$$\Sigma \frac{m_1 c^2}{(1-\mathbf{v}_1^2/c^2)^{\frac{1}{2}}} + \chi = \text{const.}, \tag{22}$$

in spite of their difference of form; so that it cannot be objected to (19) that it is not compatible with the equivalence of mass and energy.

15. *General inference.* The general inference from this investigation is that doubt must be cast on the world-wide validity of the conservation of linear momentum, in Einstein's form, for unrestricted velocities. It is seen to be always approximately true for small velocities; and it is exactly true for all velocities for a set of collinear particles in motion along their line of collinearity. But in general there appears to be no exact integral of linear or angular momentum, although there exist non-integrable relationships corresponding to them. It is strong evidence for the correctness of these relationships that they have a form which arises directly from the equations of motion, without imposing at that stage a simultaneity convention.

If surprise is felt at the non-existence of a *general* principle of conservation of momentum, it may be pointed out that the only justification for believing in such a conservation law is the law of equality of action and reaction combined with a particular form of the equations of motion. The identities that have been used in the t-dynamics are the nearest approach to a statement of Newton's third law that is possible on the t-dynamics. But even if Newton's third law is accepted, the conservation of linear momentum does not follow unless the equations of motion take the form that 'force' equals 'rate of change of momentum'. The question then arises, how is momentum to be defined? In special relativity mechanics, 'momentum' and 'mass' constitute the four-vector

$$\left(\frac{m\mathbf{v}}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}},\ \frac{mc}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}}\right),$$

and mass is not a four-scalar. Hence the so-called momentum $m\mathbf{v}/(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}$ cannot be dissected as the product of a four-scalar and a three-vector; from the point of view of Lorentz transformations, it is indissoluble, and therefore can scarcely be regarded as the natural generalization of the classical expression for momentum. But in the t-dynamics, mass is a scalar, $m\xi^{\frac{1}{2}}$, and momentum is truly the product of a scalar into a four-vector representing the velocity, namely $(\mathbf{V}/Y^{\frac{1}{2}}, c/Y^{\frac{1}{2}})$. When one transforms this to τ-measure, the counterpart of the rate of change of linear momentum, namely,

$$\frac{1}{Y^{\frac{1}{2}}}\frac{d}{dt}\left[m\xi^{\frac{1}{2}}\frac{\mathbf{V}}{Y^{\frac{1}{2}}}\right],$$

comes out to be

$$\frac{1}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}}\frac{d}{d\tau}\left[\frac{m}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}}\frac{\mathbf{v}}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}}\right],$$

and the equation of motion turns out to be the equality of this to the external force $\mathbf{\Phi}$. But when $\mathbf{\Phi}$ is associated with a potential, it has been shown in earlier papers that a term must be subtracted from $\mathbf{\Phi}$ arising from the change of mass with velocity before arriving at the gradient of the potential, according to the relation

$$\mathbf{\Phi} = -\frac{\partial\chi}{\partial\mathbf{\Pi}} + 2\frac{\mathbf{v}}{(1-\mathbf{v}^2/c^2)}\frac{d}{d\tau}\left[\frac{m}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}}\right]$$

(Scale of time, III, p. 539). The equation of motion

$$\frac{1}{(1-\mathbf{v}^2/c^2)^{\frac{1}{2}}}\frac{d}{d\tau}\frac{m\mathbf{v}}{1-\mathbf{v}^2/c^2}=\mathbf{\Phi}$$

then reduces to

$$\frac{m\,d\mathbf{v}/d\tau}{(1-\mathbf{v}^2/c^2)^{\frac{3}{2}}}=-\frac{\partial\chi}{\partial\mathbf{\Pi}},$$

in accordance with (19), but in disaccord with the conventional equation (21). The force $\mathbf{\Phi}$ cannot be asserted to obey or not to obey the law of equality of action and reaction until it is given an explicit form, and in the case of gravitational forces it is then the potential gradient $\partial\chi/\partial\mathbf{\Pi}$ which obeys this law. The weight of the train of deduction behind (19) is much more substantial than that behind (21).

It must further be pointed out that (21), the conventional form, suffers from a confusion of the two time scales, whilst (19) has an unambiguous meaning. Starting with equivalent fundamental observers whose mutual transformations of coordinates are of Lorentz type and who are using t-measure, one is led to a Lorentz-invariant dynamics of unfamiliar form, which reduces, however, to something very close to classical dynamics, and indeed very close to special relativity dynamics, on transforming the scale of time from t to τ. But (21) uses both τ-time and Lorentz transformations together. In the description of the universe in τ-measure, the fundamental observers O are relatively stationary, and these alone are equivalent to each other: another observer moving *with uniform velocity* relative to them in τ-measure is not 'equivalent' to them, and the transformations connecting such an observer's measures with the measures of fundamental observers will not be expected to be of Lorentz type. Thus, once it is seen that (21) employs τ-time and τ-measure, the arguments for the equation of motion being of that form fall to the ground. I hope to give the non-Lorentz transformations in question in a future paper.

REFERENCES

Camm, G. L. 1945 *Nature*, **155**, 754.
Eddington, A. S. 1923 *The mathematical theory of Relativity*. Cambridge.
Milne, E. A. 1936 *Proc. Roy. Soc.* A, **156**, 62.
Milne, E. A. 1937*a* *Proc. Roy. Soc.* A, **158**, 324.
Milne, E. A. 1937*b* *Proc. Roy. Soc.* A, **159**, 171, 526.
Milne, E. A. 1937*c* *Proc. Roy. Soc.* A, **160**, 1, 24.
Milne, E. A. 1938 *Proc. Roy. Soc.* A, **165**, 313, 333.
Milne, E. A. 1943 *Phil. Mag.* **34**, 73, 82, 197, 235, 246.
Schild, A. 1944 *Phys. Rev.* **66**, 340.
Whitrow, G. J. 1945 *Nature*, **156**, 365.

Kinetics of the base-catalysed bromination of diethyl malonate

By R. P. Bell, F.R.S., D. H. Everett and H. C. Longuet-Higgins

Physical Chemistry Laboratory, Oxford

(*Received* 24 *September* 1945)

The kinetics of the bromination of diethyl malonate have been investigated in aqueous solution at 25°. Under suitable conditions the rate of bromination is independent of the bromine concentration, and is determined by the rate at which the ester loses a proton to basic ions or molecules in the solution. Measurements in buffer solutions were used to obtain values for the catalytic constants of ten basic species. There is no detectable catalysis by acids, a slight increase in velocity in strongly acid solutions being traceable to hydrolysis of the ester group.

The catalytic constants of six basic anions are related to the dissociation constants of the corresponding acids by an expression of the type due to Brönsted. The ions OH^-, $H_2BO_3^-$ and OBr^- do not conform to this relation, and reasons for these discrepancies are discussed. In its general kinetic behaviour the bromination of diethyl malonate conforms to the regularities previously shown to exist in the ionization of substituted ketones (Bell 1943).

Introduction

This work represents an extension of previous measurements (Bell & Lidwell 1940) on the base-catalysed prototropy of substituted ketones. No previous kinetic work has been reported on the halogenation of diethyl malonate, though West (1924) has investigated the bromination of malonic acid in aqueous solution. He found that the rate of introduction of the first bromine atom was independent of the bromine concentration, but that the subsequent bromination of monobromomalonic acid was of the first order with respect to halogen.

Preliminary experiments showed that the iodination of diethyl malonate in aqueous solution is a reversible reaction going only partly to completion, but that bromine reacts quantitatively to give ethyl dibromomalonate. The extent of bromination at any instant can be determined by removing the excess of bromine with allyl alcohol, and then adding potassium iodide, which reacts with the dibromo-ester to give iodine. The rate of bromination was found to be unaffected by a five-fold variation in the bromine concentration, provided that this exceeded about 0·01 M. All the experiments described below were carried out with bromine concentrations of about 0·03 M. The reactions were strictly of the first order with respect to the ester, so that it can be assumed that diethyl monobromomalonate is brominated much more rapidly than the original ester. Hence the rates observed under conditions of basic catalysis represent the rates of ionization of a methylene hydrogen from the ester (cf. Bell 1941, p. 135).

EXPERIMENTAL

Materials. Commercial diethyl malonate was fractionated *in vacuo*, and the fraction boiling at 70–71°/5 mm. was used. Acetic, monochloracetic and glycollic acids were pure commercial specimens, as were the salts used. Trimethylacetic acid was prepared from pinacol hydrate as described in *Organic Syntheses* (Collective vol. 1, p. 512). The final product had b.p. 79°/20 mm., m.p. 34°. Buffer solutions were made by adding sodium hydroxide solution (free from carbonate and standardized against constant boiling hydrochloric acid) to a solution of acid which had been directly standardized against the sodium hydroxide solution.

Measurement of reaction velocity

All measurements were made at 25 ± 0·01°. In carrying out an experiment, 2 c.c. of buffer solution 0·002–0·003 M in ester were added to each of ten small glass-stoppered tubes in the thermostat. After 10 min. 1 c.c. of 0·1 M bromine in 0·1 M potassium bromide solution was added to each tube, using the type of automatic pipette previously described (Bell, Lidwell & Vaughan-Jackson 1936). After suitable periods the reaction was stopped by adding two drops of 10 % allyl alcohol solution to successive tubes.

Some difficulty was experienced in finding the right conditions for the accurate estimation of the dibromoester formed. The reaction between dibromoester and potassium iodide is always slow, and is slightly reversible under some conditions. It proceeds to completion in strongly acid solution, but under these conditions there is considerable atmospheric oxidation in the time necessary for reaction (½–1 hr.). The procedure finally adopted was to add (immediately after the allyl alcohol) 1 c.c. of molar 1 : 1 acetate buffer and a few crystals of potassium iodide. The iodine liberated was titrated with N/100 sodium thiosulphate (micro-burette), three successive titrations being carried out on each sample at 30 min. intervals, and the sum of the titres taken. The last titre never exceeded 0·02 c.c., and atmospheric oxidation was negligible. It was not found advisable to follow more than about two-thirds of each reaction, since beyond this point solid dibromoester separated out on the walls of the tube and reacted very slowly with the potassium iodide.

The titre a corresponding to complete reaction was calculated from the known ester concentration. For each reaction mixture $\log_{10}(a-x)$ was plotted against t, x being the titre at time t. This plot was invariably a straight line, showing that the reaction is of the first order with respect to ester. The velocities v in the following tables represent the slopes of these plots, the time being expressed in minutes. In the earlier measurements of Bell & Lidwell (1940) the velocities were expressed as the rate of disappearance of halogen in moles per litre per minute, referred to a ketone concentration of one mole per litre. Since each molecule of ester reacts with two molecules of bromine, the present values of v can be converted to the earlier units by multiplying by the factor $2 \times 2{\cdot}303$.

RESULTS IN BUFFER SOLUTIONS

The general plan was the same as that adopted by Bell & Lidwell (1940), i.e. in each series of experiments the anion concentration was varied at constant buffer ratio, the total ionic strength being kept constant at $\mu = 0{\cdot}1$ by the addition of sodium chloride or sodium perchlorate. The values of r in the tables represent the stoichiometric buffer ratios, and the values of c are the anion concentrations given by the amount of sodium hydroxide used in making up the buffer solutions. In obtaining c (corr.) the following corrections were applied:

(i) If the value of $[H^+]$ is not negligible compared to the anion concentration, c must be increased by an amount $[H^+]$ in order to give c (corr.). In calculating $[H^+]$ the thermodynamic dissociation constant was used, in conjunction with the value $f_\pm = 0{\cdot}8$ for the activity coefficients at $\mu = 0{\cdot}1$. The sources of the dissociation constants are given in table 4, and the value of $f_\pm$ is taken from the work of Larsson & Adell (1931) on a variety of carboxylic acids.

(ii) The anion concentration is reduced slightly by the acid produced during the reaction. Each molecule of ester brominated produces two hydrogen ions, and approximately the first two-thirds of each reaction was followed: hence two-thirds of the initial ester concentration was subtracted from c to obtain the true average anion concentration. In the glycollate solutions an additional correction was necessary for the production of hydrogen ions in the slow oxidation of the glycollate by bromine. This reaction was studied in separate experiments, and contributed only a small correction in the acid buffer solutions used.

(iii) In the more alkaline buffer solutions the bromine was hydrolysed sufficiently to displace the anion concentration to a small extent. This hydrolysis takes place according to the equation $Br_2 + H_2O \rightleftharpoons H^+ + Br^- + HOBr$, and the best value for the constant $f_\pm^2[H^+][Br^-][HOBr]/[Br_2]$ at 25° is $5{\cdot}8 \times 10^{-9}$ (Liebhafsky 1935). In computing $[Br_2]$ and $[Br^-]$ allowance was made for the equilibrium $Br_2 + Br^- \rightleftharpoons Br_3^-$, the constant $[Br_2][Br^-]/[Br_3^-]$ being taken as $0{\cdot}065$ (Jakowkin 1896). A correction equal to [HOBr] was then subtracted from the anion concentration, since one hydrogen ion is produced for each molecule of hypobromous acid.

(iv) Experiments in solutions of bromine and bromide (*vide infra*) showed that the hypobromite ion makes a contribution to the velocity of approximately $v(OBr^-) = 1150[HOBr]^2[Br^-]/[Br_2]$. This is negligible in most of the other experiments, but is appreciable in the more alkaline acetate and trimethylacetate buffers. For these experiments $v(OBr^-)$ has been subtracted from v (obs.) to give v (corr.).

All the above corrections are small, and it is legitimate to apply them independently of each other. They were omitted in the earlier work of Bell & Lidwell (1940), but their insertion would not materially affect most of the conclusions reached. However, the value previously given for hydroxyl-ion catalysis in the bromination of *as*-dichloracetone and the analogous value reported provisionally for diethyl malonate (Bell 1943) must be rejected, as they depend on inaccurate extrapolations in buffer solutions where the corrections for bromine hydrolysis are considerable.

The data obtained for buffer solutions are given in table 1. In the case of trimethylacetate, acetate, glycollate and monochloroacetate buffers the values of v (calc.) are obtained from the relation $v = 0{\cdot}00062 + k_b c$, where 0·00062 represents the rate due to water catalysis, and k_b is the catalytic constant of the anion. The value 0·00062 is derived from the results in phosphate buffers, and is confirmed by the measurements in acid solutions described below. The success of this simple equation in acetate buffers of ratios varying from 0·16 to 2·2 demonstrates that there is no detectable acid catalysis, and this conclusion is confirmed also by the measurements in acid solutions. The simple form of the equation also shows that there is no detectable catalysis by hydroxyl ions in the buffers in question: this is confirmed by the measurements in bromine solutions described in the next section.

In the phosphate buffers the slope of the plot of v against $[H_2PO_4^-]$ was found to decrease appreciably as the ratio $[H_3PO_4]/[H_2PO_4^-]$ increases. This is because the small concentrations of $HPO_4^=$ ions present are exerting an appreciable catalytic effect, and in fact the data are sufficiently accurate to yield approximate values for the catalytic constants of both these phosphate ions. The values of $[H_3PO_4]$ (corr.) and $[H_2PO_4^-]$ (corr.) were calculated from the stoichiometric buffer composition, using the value $[H^+][H_2PO_4^-]/[H_3PO_4] = 8{\cdot}71 \times 10^{-3}$ given by Bjerrum & Unmack (1929) for $\mu = 0{\cdot}1$ at 25°, and applying correction (ii) above. The values of v (calc.) in the table are then given in terms of the corrected concentrations by the expression

$$10^4 v = 6{\cdot}2 + 66[H_2PO_4^-] + 7{\cdot}2[H_2PO_4^-]^2/[H_3PO_4].$$

In this expression, the first term represents water catalysis and the second term catalysis by the ion $H_2PO_4^-$, with a catalytic constant of $6{\cdot}6 \times 10^{-3}$. The concentration of the ion $HPO_4^=$ in these solutions is given by $[HPO_4^=] = (K_c''/K_c')[H_2PO_4^-]^2/[H_3PO_4]$, where K_c' and K_c'' are the first and second concentration dissociation constants of phosphoric acid. The third term thus represents catalysis by the ion $HPO_4^=$, and its catalytic constant is $k(HPO_4^=) = 7{\cdot}2 \times 10^{-4} K_c'/K_c''$. Taking $K_c' = 8{\cdot}71 \times 10^{-3}$, $K_c'' = 1{\cdot}79 \times 10^{-7}$ for $\mu = 0{\cdot}1$ at 25° (Bjerrum & Unmack 1929), this gives $k(HPO_4^=) = 35$. The values obtained for the catalytic constants of the two phosphate ions are rather sensitive to the values assumed for the dissociation constants of phosphoric acid in 0·1 N salt solutions. They are therefore subject to a much greater uncertainty than the values for carboxylate ions: thus $k(H_2PO_4^-)$ may well be in error by 10 %, and $k(HPO_4^=)$ by considerably more.

The experiments with boric acid in an acetate buffer were designed to detect catalysis by the borate ion. The results show that the reaction velocity is not measurably affected by the addition of boric acid. By using the value for the dissociation constant of boric acid given by Owen (1934) it is found that the concentration of borate ions in these solutions is approximately 10^{-4} times that of boric acid. Since the addition of 0·2 M boric acid changes v by less than 1×10^{-3}, a reasonable upper limit for the catalytic constant of the ion $H_2BO_3^-$ is $0{\cdot}001/(2 \times 10^{-5}) = 50$.

TABLE 1. MEASUREMENTS IN BUFFER SOLUTIONS

Trimethylacetate: $k_b = 0\cdot97$, $r = 1\cdot25$

$10^4 c$	664	332	166	100
$10^4 c$ (corr.)	654	322	156	90
$10^4 v$: obs.	652	330	155	92
corr.	650	328	153	90
calc.	641	318	157	93

Acetate: $k_b = 0\cdot59$

	$r = 0\cdot164$			$r = 0\cdot600$			
$10^4 c$	728	549	366	663	498	332	166
$10^4 c$ (corr.)	695	518	336	647	482	316	150
$10^4 v$: obs.	448	339	224	396	296	197	102
corr.	422	320	217	390	290	191	96
calc.	428	326	223	388	291	193	95

	$r = 0\cdot99$				$r = 2\cdot20$			
$10^4 c$	728	549	366	183	663	498	332	166
$10^4 c$ (corr.)	704	527	344	161	653	488	322	156
$10^4 v$: obs.	428	315	211	108	376	285	195	102
corr.	426	313	209	106	—	—	—	—
calc.	421	318	210	102	391	294	196	98

Glycollate: $k_b = 9\cdot25 \times 10^{-2}$, $r = 4\cdot96$

$10^4 c$	681	513	341	171
$10^4 c$ (corr.)	666	503	341	167
$10^4 v$: obs.	65·3	51·6	38·4	22·4
calc.	67·7	52·7	37·7	21·6

Monochloroacetate: $k_b = 1\cdot90 \times 10^{-2}$, $r = 0\cdot555$

$10^4 c$	645	483	321	160
$10^4 c$ (corr.)	639	477	316	154
$10^4 v$: obs.	17·8	15·1	12·2	9·4
calc.	18·3	15·3	12·2	9·1

Phosphate: $10^4 v$ (calc.) $= 6\cdot2 + 66[H_2PO_4^-] + 7\cdot2[H_2PO_4^-]^2/[H_3PO_4]$

	$r = 0\cdot246$				$r = 0\cdot412$			
$10^4[H_2PO_4^-]$	680	510	340	170	678	509	339	170
$10^4[H_2PO_4^-]$ (corr.)	682	510	343	173	692	522	350	176
$10^4[H_3PO_4]$ (corr.)	166	126	81	39	265	197	129	64
$10^4 v$: obs.	12·4	11·0	9·5	7·9	12·1	10·5	9·2	7·8
calc.	12·7	11·0	9·5	7·9	12·1	10·7	9·2	7·7

	$r = 0\cdot975$			
$10^4[H_2PO_4^-]$	680	509	340	170
$10^4[H_2PO_4^-]$ (corr.)	726	553	376	197
$10^4[H_3PO_4]$ (corr.)	618	452	296	139
$10^4 v$: obs.	11·6	10·3	9·1	7·8
calc.	11·6	10·3	9·1	7·7

Acetate + boric acid: Mean $r = 0\cdot32$, mean $[CH_3COO^-] = 0\cdot02$

$10^4[H_3BO_3]$	2430	1350	675	0
$10^4 v$ (obs.)	120	118	115	117

RESULTS IN BROMINE SOLUTIONS

It was hoped that these experiments would lead to a value for the catalytic constant of the hydroxyl ion. In a solution made by adding bromine to sodium hydroxide (which may or may not contain added bromide ion initially), the concentrations of hydrogen and hydroxyl ions are controlled by the equilibrium $Br_2 + H_2O \rightleftharpoons H^+ + Br^- + HOBr$. Such a solution thus constitutes a kind of triple buffer mixture, in which pH can be controlled in the range pH = 5–7 by varying the concentrations of Br_2, Br^- and HOBr. Moreover, the only basic species present in these solutions (apart from the hydroxyl ion and the water molecule) is the ion OBr^-, which by analogy with the borate ion might be expected to make only a small contribution to the reaction velocity. Conditions should therefore be specially favourable for measuring hydroxyl-ion catalysis.

In order to minimize the loss of bromine vapour, a different experimental technique was used in these measurements. 50 c.c. of reaction mixture was made up in a cylindrical reaction vessel about 3 cm. in diameter, and samples for analysis were ejected from the bottom of the reaction vessel by pressure. The total oxidizing power of the solution was determined at the beginning and at the end of the reaction by adding a sample to potassium iodide in an acetate buffer and titrating the iodine liberated. (If nothing is lost from the solution these two titres should agree, since every two moles of HOBr or Br_2 are replaced by one mole of $CBr_2(COOEt)_2$ having the same oxidizing power.) From this titre, together with the concentrations of sodium hydroxide and sodium bromide used in making up the solution, it was possible to calculate the initial concentrations of Br_2, Br^- and HOBr, by using the equilibrium constants $[Br_3^-] = 16[Br^-][Br_2]$, $[Br_5^-] = 40[Br^-][Br_2]^2$ (Jakowkin 1896; Jones & Baeckström 1934). The ionic strength of all solutions was made up to 0·1 by adding the appropriate amounts of sodium chloride or perchlorate. The concentrations of Br_2, Br^- and HOBr given in table 2 are mean concentrations, corrected for the acid produced in the reaction of the ester. It was found necessary to modify the method of estimating the amount of dibromoester formed, since a mixture of triple buffer with allyl alcohol and acetate buffer was found to liberate iodine slowly from potassium iodide. This behaviour is probably due to the presence of small quantities of bromate. To avoid this difficulty, samples of reaction mixture were added to weighed tubes containing N hydrochloric acid, which effectively stops the reaction and converts any bromate into bromine. Allyl alcohol solution was then added, and the titration completed in the usual way.

TABLE 2. MEASUREMENTS IN BROMINE SOLUTIONS

$10^4[HOBr]$	180	180	180	87	87	87	44	44	44	44
$10^4[Br_2]$	420	413	270	350	221	148	233	211	189	165
$10^4[Br^-]$	104	147	198	92	551	214	162	683	39	106
$10^3 v$: obs.	109	156	264	29	202	107	20	72	7	11
calc.	97	134	287	26	218	130	17	73	5	15

The results are given in table 2. They show rather large erratic variations, which may be due to the formation of some bromate in the hypobromite solutions, or to the imperfect exclusion of carbon dioxide. However, there are no systematic deviations between the observed velocities and those calculated from the equation

$$v = 0{\cdot}00062 + 1150[\mathrm{HOBr}]^2[\mathrm{Br}^-]/[\mathrm{Br}_2].$$

The first term represents the water catalysis, and the second is proportional to $[\mathrm{HOBr}]/[\mathrm{H}^+]$, i.e. to $[\mathrm{OBr}^-]$. This shows that (contrary to expectation) the hypobromite ion is the chief catalyst in these solutions, which are thus not suitable for observing hydroxyl-ion catalysis. It is easily seen that

$$[\mathrm{OBr}^-] = K(\mathrm{HOBr})[\mathrm{HOBr}]^2[\mathrm{Br}^-]/K_h[\mathrm{Br}_2],$$

where $K(\mathrm{HOBr})$ is the thermodynamic dissociation constant of hypobromous acid ($2{\cdot}1 \times 10^{-9}$, Shilov & Gladtchikova 1938) and K_h is the hydrolysis constant of bromine ($5{\cdot}8 \times 10^{-9}$, Liebhafsky 1935). The catalytic constant of the hypobromite ion is therefore given by $k(\mathrm{OBr}^-) = 1150 \times 5{\cdot}8/2{\cdot}1 = 2{\cdot}8 \times 10^{+3}$. If it is assumed that hydroxyl-ion catalysis contributes less than one-sixth of the total velocity in the experiments with the lowest concentration of hypobromous acid, then the upper limit for $k(\mathrm{OH}^-)$ works out at $5 \times 10^{+5}$.

RESULTS IN SOLUTIONS OF STRONG ACIDS

In very dilute acid solutions the rate of bromination is approximately constant at $v = 6{\cdot}2 \times 10^{-4}$, agreeing with the value for the water rate derived from the experiments in phosphate buffers. However, when the acid concentration is increased, the measured rate increases slowly, reaching $v = 7{\cdot}7 \times 10^{-4}$ in $0{\cdot}25$N hydrochloric acid. This increase might be interpreted as catalysis by hydrogen ions, but it is first necessary to consider the effect of slow hydrolysis of the ester group by the acid. The total reaction scheme would be

$$\begin{tikzcd}
\text{diester} \arrow[r, "k_1"'] \arrow[d, "j_1"'] & \text{monoester} \arrow[r, "k_3"'] \arrow[d, "j_2"'] & \text{malonic acid} \arrow[d, "j_3"'] \\
\text{dibromodiester} \arrow[r, "k_2"'] & \text{dibromomonoester} \arrow[r, "k_4"'] & \text{dibromoacid}
\end{tikzcd}$$

In this scheme $j_1 \simeq 6 \times 10^{-4}$, and the measurements of Skrabal & Mrazek (1918) show that in our present units $k_1 = 8{\cdot}9 \times 10^{-4}[\mathrm{H}^+]$, $k_3 = 4{\cdot}5 \times 10^{-4}[\mathrm{H}^+]$. Therefore, if $[\mathrm{H}^+] > 0{\cdot}01$, the rate of hydrolysis of the diester is not negligible compared with its rate of bromination, though the subsequent hydrolysis of the monoester can be neglected for any moderate acid concentrations. In order to correct for the first-stage hydrolysis it was necessary to make a rough investigation of the bromination of monoethyl malonate. This substance can be prepared in solution by the alkaline hydrolysis of diethyl malonate, which is much faster than the subsequent alkaline hydrolysis of the monoester. One equivalent of sodium hydroxide solution was added to a M/100 solution of the diester, and the mixture left for 30 min. A slight excess of hydrochloric acid was then added, and the resulting solution used for a bromination experiment in the usual way. The results of a typical experiment are given in table 3. It will be seen that the dibromomonoester first formed is progressively converted into some product which does not liberate iodine from potassium iodide. Similar behaviour has been reported in the bromination of acetoacetic ester (Pedersen 1933) and of acetylacetone (Bell & Lidwell 1940), and in the present instance is probably due to slow decomposition according to the equation

$$\mathrm{COOH.CBr_2.COOEt \rightarrow CHBr_2.COOEt + CO_2}.$$

TABLE 3. BROMINATION OF MONOETHYL MALONATE

Initial $[\mathrm{H^+}] = 2{\cdot}1 \times 10^{-3}$, initial [ester] $= 2{\cdot}77 \times 10^{-3}$

t (min.)	0	1	$2\frac{1}{2}$	4	8	16	31	65	93	145	196	308
10^5 [dibromoester]	0	6	13	20	30	51	97	123	117	94	62	23

The course of the change can be satisfactorily represented by two consecutive first-order reactions, the formation of dibromoester having $k \simeq 5 \times 10^{-3}$, $\log_{10}$/min., and its destruction $k \simeq 4 \times 10^{-3}$. The reaction scheme must thus be modified to

$$\begin{tikzcd}
\underset{(a)}{\text{diester}} \arrow[r, "k_1"] \arrow[d, "j_1"] & \underset{(b)}{\text{monoester}} \arrow[d, "j_2"] & \\
\underset{(a')}{\text{dibromodiester}} \arrow[r, "k_2"'] & \underset{(b')}{\text{dibromomonoester}} \arrow[r, "m"'] & \text{inactive product}
\end{tikzcd}$$

in which a, b, etc., denote concentrations, and we have approximately $10^4 j_1 = 6{\cdot}2$, $10^4 k_1 = 8{\cdot}9[\mathrm{H^+}]$, $10^4 j_2 = 50$, $10^4 m = 40$, k_2 being unknown. If a_0 is the initial concentration of the diester, the following expressions give the values of a' and b' after time t:

$$a' = a_0\{10^{-k_1 t} - 10^{-(j_1+k_1)t}\},$$

$$b' = \frac{j_2 k_1 a_0}{j_2 - k_1 - j_1}\left\{\frac{10^{-(k_1+j_1)t} - 10^{-mt}}{m - k_1 - j_1} - \frac{10^{-j_2 t} - 10^{-mt}}{m - j_2}\right\} + \frac{j_1 k_2 a_0}{k_2 - k_1 - j_1}\left\{\frac{10^{-(k_1+j_1)t} - 10^{-mt}}{m - k_1 - j_1} - \frac{10^{-k_2 t} - 10^{-mt}}{m - k_2}\right\}.$$

By analogy with the known effects of halogen substitution on the acid hydrolysis of simple carboxylic esters, it is probably safe to assume that k_2 is not greater than k_1. If this is so, the above expressions show that when $[\mathrm{H^+}] > 0{\cdot}1$, b' is negligible compared with a', i.e. the observed titres are almost entirely due to dibromodiester. These conditions are satisfied by the first four experiments in table 4. In these experiments it was therefore possible to derive a' directly from the titre, and to obtain an accurate value for $j_1 + k_1$ from the slope of the plot of $\log_{10}\{a_0 . 10^{-k_1 t} - a'\}$ against the time. In the last experiment with $[\mathrm{H^+}] = 0{\cdot}25$, values of b' were calculated from the above expression using the approximate value for j_1, and making the arbitrary assumption that $k_2 = k_1$. These values of b' were subtracted from the observed titres to obtain a', and j_1+k_1 evaluated as before. Owing to the arbitrary assumption made about k_2 this last experiment cannot be given much weight. However, the data show clearly that the apparent increase in rate with increasing acid concentration can reasonably be accounted for by the complications due to hydrolysis, and hence that there is no evidence for catalysis of the bromination reaction by hydrogen ions. The mean value for j_1 agrees well with the value $6{\cdot}2\times10^{-4}$ derived for water catalysis from the measurements in phosphate buffers.

TABLE 4. MEASUREMENTS IN ACID SOLUTIONS

$[H^+]$ (initial)	0·001	0·01	0·02	0·05	0·25
$10^4(j_1+k_1)$	6·2	6·3	6·5	6·8	(7·7)
$10^4 j_1$	6·2	6·2	6·3	6·3	(5·5)

DISCUSSION OF RESULTS

The above measurements demonstrate that the rate of bromination of diethyl malonate is determined by the rate at which the methylene group loses a proton to a basic ion or molecule in the solution. The catalytic constants of the different basic species are collected in table 5, together with the thermodynamic dissociation constants K_{HA} of the corresponding acids. The calculated catalytic constants are derived from the equation $k_b = 1{\cdot}05\times10^{-4}(1/K_{HA})^{0{\cdot}79}$. There is good agreement between observed and calculated values for the anions of the five acids with dissociation constants between 10^{-2} and 10^{-6}. It is remarkable that this agreement should include the primary phosphate ion, since in comparing this ion with carboxylate ions a statistical correction of $3^{0{\cdot}79}$ should properly be applied, and is, in fact, necessary to obtain agreement in the catalysed decomposition of nitramide (cf. Bell 1941, pp. 82–87). No special significance can be attached to the small discrepancies for the catalysts $HPO_4^=$ and H_2O, since these differ in charge type from the other catalysts. In the case of $H_2BO_3^-$ and OH^- the upper limits obtained experimentally for the catalytic constants are smaller by several powers of ten than the calculated values. It is likely that these discrepancies represent genuine deviations from the simple Brönsted relation, which would not be expected to be valid over an unlimited range of basic strengths. The hydroxyl ion is commonly much less active as a catalyst than would be predicted by extrapolation from other basic anions (cf. Bell 1941, p. 92) and a similar negative deviation has been found for catalysis by the phenate ion ($K_{HA} = 1{\cdot}3\times10^{-10}$) in the decomposition of nitramide (Tong & Olson 1941). It has been pointed out (Bell 1943) that, on account of the association of water in the liquid state, the figure $55{\cdot}5/K_w$ gives an exaggerated view of the basic strength of the hydroxyl ion, and an attempt to allow for the effect of association led to the corrected value $0{\cdot}13/K_w$ as a lower limit for this basic strength. However, in the present reaction even this value leads to a calculated catalytic constant for OH^- five times as great as the upper limit found experimentally.

TABLE 5. SUMMARY OF CATALYTIC CONSTANTS

catalyst	K_{HA}	k_b (obs.)	k_b (calc.)
OH^-	$1\cdot8 \times 10^{-16}$ (Harned & Hamer 1933)	$< 5 \times 10^{+5}$	$3\cdot0 \times 10^{+8}$
$H_2BO_3^-$	$5\cdot8 \times 10^{-10}$ (Owen 1934)	$< 5 \times 10^{+1}$	$2\cdot1 \times 10^{+3}$
OBr^-	$2\cdot1 \times 10^{-9}$ (Shilov & Gladtchikova 1938)	$2\cdot8 \times 10^{+3}$	$7\cdot9 \times 10^{+2}$
$HPO_4^=$	$6\cdot2 \times 10^{-8}$ (Nims 1933)	$3\cdot5 \times 10^{+1}$	$5\cdot3 \times 10^{+1}$
$Me_3C.COO^-$	$9\cdot1 \times 10^{-6}$ (Dippy 1938)	$9\cdot7 \times 10^{-1}$	$10\cdot0 \times 10^{-1}$
CH_3COO^-	$1\cdot75 \times 10^{-5}$ (Harned & Ehlers 1933)	$5\cdot9 \times 10^{-1}$	$6\cdot0 \times 10^{-1}$
$CH_2OH.COO^-$	$1\cdot54 \times 10^{-4}$ (Nims 1936)	$9\cdot3 \times 10^{-2}$	$10\cdot7 \times 10^{-2}$
$CH_2Cl.COO^-$	$1\cdot38 \times 10^{-3}$ (Wright 1934)	$1\cdot9 \times 10^{-2}$	$1\cdot9 \times 10^{-2}$
$H_2PO_4^-$	$7\cdot52 \times 10^{-3}$ (Nims 1934)	$6\cdot6 \times 10^{-3}$	$5\cdot0 \times 10^{-3}$
H_2O	55·5	$1\cdot1 \times 10^{-5}$	$4\cdot4 \times 10^{-6}$

On the other hand, the observed catalytic constant for the hypobromite ion is considerably *greater* than the calculated value, though the dissociation constants of HOBr and H_3BO_3 are similar. This probably indicates that the rate-determining step is not (as with other anions) $CH_2(COOEt)_2 + OBr^- \rightarrow \bar{C}H(COOEt)_2 + HOBr$ followed by rapid bromination of the ester anion. Instead, the hypobromite ion may brominate the ester molecule by a single-stage reaction, e.g.

$$CH_2(COOEt)_2 + OBr^- \rightarrow CHBr(COOEt)_2 + OH^-.$$

If this is the case the observed velocity in presence of hypobromite ions would not be comparable with the values for other anions.

Malonic ester can be considered as a member of the series of substituted ketones whose rates of ionization have previously been investigated (Bell & Lidwell 1940; Bell 1943). The exponent of the Brönsted relation is $\alpha = 0\cdot79$, and the rate of ionization as previously defined is $\log_{10} R = -0\cdot46$. These values fit well into the series between *as*-dichloroacetone ($\alpha = 0\cdot82, \log_{10} R = -2\cdot00$) and ethyl acetoacetate ($\alpha = 0\cdot59, \log_{10} R = +1\cdot02$).

One of us (D. H. E.) wishes to acknowledge his indebtedness to the Trustees of the Ramsay Memorial Fund for a Fellowship which enabled him to take part in this work.

REFERENCES

Bell, R. P. 1941 *Acid-base catalysis*. Oxford: Clarendon Press.
Bell, R. P. 1943 *Trans. Faraday Soc.* **39**, 253.
Bell, R. P. & Lidwell, O. M. 1940 *Proc. Roy. Soc.* A, **176**, 88.
Bell, R. P., Lidwell, O. M. & Vaughan-Jackson, M. W. 1936 *J. Chem. Soc.* p. 1792.
Bjerrum, N. & Unmack, A. 1929 *Math.-fys. Medd.* **9**, 1.
Dippy, J. F. J. 1938 *J. Chem. Soc.* p. 1222.
Harned, H. S. & Ehlers, R. W. 1933 *J. Amer. Chem. Soc.* **55**, 652.
Harned, H. S. & Hamer, W. J. 1933 *J. Amer. Chem. Soc.* **55**, 2194.
Jakowkin, A. A. 1896 *Z. Phys. Chem.* **20**, 19.
Jones, G. & Baeckström, S. 1934 *J. Amer. Chem. Soc.* **56**, 1517.
Larsson & Adell 1931 *Z. phys. Chem.* A, **156**, 352, 381.
Liebhafsky, H. A. 1935 *Chem. Rev.* **17**, 89.
Nims, L. F. 1933 *J. Amer. Chem. Soc.* **55**, 1946.

Nims, L. F. 1934 *J. Amer. Chem. Soc.* **56**, 1110.
Nims, L. F. 1936 *J. Amer. Chem. Soc.* **58**, 987.
Owen, B. B. 1934 *J. Amer. Chem. Soc.* **56**, 1695.
Pedersen, K. J. 1933 *J. Phys. Chem.* **37**, 751.
Shilov, E. A. & Gladtchikova, J. N. 1938 *J. Amer. Chem. Soc.* **60**, 490.
Skrabal, A. & Mrazek, D. 1918 *S.B. Akad. Wiss. Wien*, **127**, 301.
Tong, L. K. J. & Olson, A. R. 1941 *J. Amer. Chem. Soc.* **63**, 3406.
West, R. W. 1924 *J. Chem. Soc.* p. 1277.
Wright, D. D. 1934 *J. Amer. Chem. Soc.* **56**, 314.

An invariant form for the prior probability in estimation problems

By Harold Jeffreys, F.R.S.

(*Received* 23 *November* 1945)

It is shown that a certain differential form depending on the values of the parameters in a law of chance is invariant for all transformations of the parameters when the law is differentiable with regard to all parameters. For laws containing a location and a scale parameter a form with a somewhat restricted type of invariance is found even when the law is not everywhere differentiable with regard to the parameters. This form has the properties required to give a general rule for stating the prior probability in a large class of estimation problems.

1. *The consistency of induction.* The consistency of pure mathematics has never been in serious doubt, but work mostly done in the present century has shown that a definite proof that pure mathematics is consistent is a matter of considerable difficulty. Nevertheless, both the creators and the users of pure mathematics have proceeded on the hypothesis that consistency at least cannot be disproved. Suitable axioms for the theory of real numbers have been stated, and if those of any extension can be formulated in terms of those of real numbers that has been held to be sufficient justification.

The usual attitude towards generalization from experience is quite different. It is taken for granted that it is possible, but when it comes to actual application there is seldom any attempt to state formal rules, and different workers proceed by methods that lead to different conclusions from the same data. Each is sure that there is only one right method, but they disagree about the fundamental principles.

The object of my *Theory of Probability* (1939) is to construct a system whose most important properties should be that (1) on the same evidence the probabilities of a given set of propositions shall have a unique order, (2) the axioms of the system shall not deny any law capable of being stated as a prediction concerning observations, (3) with sufficient observational evidence any such law can be established with a probability approaching certainty. Property (1) is the assertion that a consistent scheme is possible, but merely making this assertion does not guarantee that a given scheme is consistent. It does make it possible to derive theorems by equating probabilities found in different ways; and if, in spite of all efforts, probabilities found in different ways are different it makes it impossible to accept the situation as satisfactory. A theorem that has been established suffices to establish consistency over a large part of the subject, namely, that if q_1 and q_2 are two alternative hypotheses, p_1 and p_2 two sets of observational data, then the posterior probabilities of q_1 and q_2, given the same prior probabilities, are in the same ratio whether p_1 is taken into account and then p_2, the order reversed, or both allowed for at once. The proof depends on supposing that the product rule for probabilities holds for the likelihoods, but this has never been questioned. It follows at once from the definition if the probability of p_2, given $p_1 q_r H$, is the same as the probability of p_2, given $q_r H$ alone; that is, if the likelihoods are chances.* It is not obviously true in general, but it does follow that, if the product rule for probabilities is consistent for likelihoods, no inconsistency can arise through the principle of inverse probability.

The question then arises, can any general rule be laid down for assessing the prior probability itself? In existing theory there is a principle that specifies it for any quantity such as a standard error, where it might be equally reasonable to regard some power of the parameter as fundamental; the only rule that is invariant (except for an irrelevant constant) for such changes is the $d\sigma/\sigma$ rule. In other cases uniform distribution is used in estimation problems, and various more or less satisfactory rules are adopted in significance tests, but no general rule is available, and the absence of one makes further extensions difficult. For errors satisfying a Pearson law of Type VII, for instance, it is natural to use the $d\sigma/\sigma$ rule for the scale parameter and uniform distribution for the location parameter, but there is no definite rule to say what should be taken for the index—or, to put it in another way, what function of the index should be taken to have uniform prior probability distribution.

An attempt to give a general rule will require a new set of conditions. As long as each type of problem is treated separately, no problem of consistency arises; the prior probability in one type is irrelevant to that in another, and there is no more to be said. But it may turn out (1) that no general rule can be used consistently, (2) that there may be a choice of several, or that (3) a general rule may be self-consistent but contrary to the fundamental principles that we are trying to express in our formalism, or (4) conflict with other less fundamental principles that we are not prepared to abandon lightly. A satisfactory rule must, in particular, provide for the distinction between problems of estimation and significance.

2. *Invariants expressing the difference between two distributions of chance.* For simplicity take the laws as referring to one variable x; let the respective chances that x is not greater than an assigned value be P, P'. Then consider the integrals

$$I_1 = \int (\sqrt{dP'} - \sqrt{dP})^2, \quad I_2 = \int \log \frac{dP'}{dP}\, d(P'-P), \tag{1}$$

---

* This is used in the proof of Theorem 9 (1939, p. 24) but has escaped explicit statement.

defined in the Stieltjes manner, by taking δP, $\delta P'$ for the same interval δx, forming the approximating sums, and letting the intervals of x tend to zero. It is obvious from the form of these integrals, since they depend only on the distribution of P, P', that they are invariant for all transformations of the parameters in the laws and also for all non-singular transformations of x, the range of integration being over the range permitted to x. Further, they are both positive definite. They therefore provide measures of the discrepancy between the two laws.

The greatest possible discrepancies between the laws will be when $dP' = 0$ in all ranges where P varies, and conversely. In that case $I_1 = 2$ and I_2 is infinite. The same extreme values occur if P is continuous as a function of x but P' varies only at isolated points.

These are not the only invariants with similar properties; any integral of the form

$$\int |(dP')^m - (dP)^m|^{1/m} \tag{2}$$

would behave similarly; but I_1 and I_2 are apparently the only ones that are ordinarily of the second order in the differences of the parameters in the laws when these differences are small.

If the range of x is subdivided and $p_r = \delta P_r, p_r = \delta P'_r$ for the rth interval, and if p'_r is differentiable with regard to all parameters in the law, then to the first order

$$\sqrt{(\delta P'_r)} - \sqrt{(\delta P_r)} = \sqrt{p'_r} - \sqrt{p_r} = \frac{1}{2\sqrt{p_r}}(p'_r - p_r), \tag{3}$$

$$\log\frac{\delta P'_r}{\delta P_r} = \frac{1}{p_r}(p'_r - p_r), \tag{4}$$

and to the second order

$$I_1 = \lim \Sigma \frac{1}{4}\frac{(p'_r - p_r)^2}{p_r}, \quad I_2 = \lim \Sigma \frac{(p'_r - p_r)^2}{p_r}. \tag{5}$$

For laws referring to more than one variable only formal changes are needed. The limit taken corresponds to $(\delta x_r)_{\max} \to 0$. Now if the P_r law depends on parameters α_i $(i = 1 \text{ to } m)$, and P'_r on $\alpha_i + \Delta\alpha_i$,

$$I_2 = \lim \Sigma \frac{1}{p_r}\left(\frac{\partial p'_r}{\partial \alpha_i}\Delta\alpha_i\right)\left(\frac{\partial p'_r}{\partial \alpha_k}\Delta\alpha_k\right)$$

$$= g_{ik}\Delta\alpha_i\Delta\alpha_k, \tag{6}$$

where
$$g_{ik} = \lim_{\delta x_r \to 0} \Sigma \frac{1}{p_r}\frac{\partial p'_r}{\partial \alpha_i}\frac{\partial p'_r}{\partial \alpha_k}. \tag{7}$$

Hence for small variations of the parameters I_2, and therefore I_1, have the form of the square of an element of distance in curvilinear co-ordinates. If we transform to any other parameters α'_j, I_2 is unaltered, and

$$I_2 = g'_{jl}\Delta\alpha'_j\Delta\alpha'_l, \tag{8}$$

$$g'_{jl} = g_{ik}\frac{\partial\alpha_i}{\partial\alpha'_j}\frac{\partial\alpha_k}{\partial\alpha'_l}, \tag{9}$$

$$\|g'_{jl}\| = \|g_{ik}\|\left\|\frac{\partial\alpha_i}{\partial\alpha'_j}\right\|^2. \tag{10}$$

But

$$d\alpha_1 d\alpha_2 \dots d\alpha_m = \left\|\frac{\partial\alpha_i}{\partial\alpha'_j}\right\| d\alpha'_1 \dots d\alpha'_m$$

$$= \left\{\frac{\|g'_{jl}\|}{\|g_{ik}\|}\right\}^{\frac{1}{2}} d\alpha'_1 \dots d\alpha'_m. \tag{11}$$

Hence

$$\|g_{ik}\|^{\frac{1}{2}} d\alpha_1 d\alpha_2 \dots d\alpha_m \text{ is invariant,} \tag{12}$$

and if the prior probability density of the parameters is taken as proportional to $\|g_{ik}\|^{\frac{1}{2}}$, the prior probability over any region will be invariant for all ways of choosing the parameters.

It remains to be seen whether such a choice agrees with those that have been used hitherto. Take first the normal law of error. Here

$$p_r \doteqdot \frac{1}{\sqrt{(2\pi)}\,\sigma}\exp\left\{-\frac{(x_r-\alpha)^2}{2\sigma^2}\right\}\delta x_r \tag{13}$$

and, exactly,

$$I_1 = \int_{-\infty}^{\infty}\frac{1}{\sqrt{(2\pi)}}\left[\frac{1}{\sqrt{\sigma'}}\exp\left\{-\frac{(x-\alpha')^2}{4\sigma'^2}\right\}-\frac{1}{\sqrt{\sigma}}\exp\left\{-\frac{(x-\alpha)^2}{4\sigma^2}\right\}\right]^2 dx$$

$$= 2-\int_{-\infty}^{\infty}\frac{2}{\sqrt{(2\pi)}\sqrt{(\sigma\sigma')}}\exp\left\{-\frac{(x-\alpha')^2}{4\sigma'^2}-\frac{(x-\alpha)^2}{4\sigma^2}\right\}dx$$

$$= 2\left[1-\frac{\sqrt{2}}{\sqrt{(\sigma/\sigma'+\sigma'/\sigma)}}\exp\left\{-\frac{(\alpha-\alpha')^2}{4(\sigma^2+\sigma'^2)}\right\}\right]. \tag{14}$$

Writing $\sigma = \sigma_0 e^{-\zeta}$, $\sigma' = \sigma_0 e^{\zeta}$, we have

$$I_1 = 2\left[1-\operatorname{sech}^{\frac{1}{2}} 2\zeta\exp\left(-\frac{(\alpha-\alpha')^2}{8\sigma_0^2\cosh 2\zeta}\right)\right], \quad \doteqdot 2\zeta^2+\frac{(\alpha-\alpha')^2}{4\sigma_0^2}, \tag{15}$$

when ζ is small.

Again,

$$I_2 = \int_{-\infty}^{\infty}\frac{1}{\sqrt{(2\pi)}}\left\{\log\frac{\sigma}{\sigma'}+\frac{(x-\alpha)^2}{2\sigma^2}-\frac{(x-\alpha')^2}{2\sigma'^2}\right\}\left[\frac{1}{\sigma'}\exp\left\{-\frac{(x-\alpha')^2}{2\sigma'^2}\right\}-\frac{1}{\sigma}\exp\left\{-\frac{(x-\alpha)^2}{2\sigma^2}\right\}\right]dx$$

$$= -1+\frac{\sigma'^2}{2\sigma^2}+\frac{\sigma^2}{2\sigma'^2}+\frac{(\alpha-\alpha')^2(\sigma^2+\sigma'^2)}{2\sigma^2\sigma'^2} \tag{16}$$

$$\doteqdot 8\zeta^2+\frac{(\alpha-\alpha')^2}{\sigma_0^2} \tag{17}$$

$$\doteqdot 2\left(\frac{\delta\sigma}{\sigma}\right)^2+\frac{(\delta\alpha)^2}{\sigma^2}.$$

Three cases arise. If α is fixed, $g_{\sigma\sigma} = 2/\sigma^2$, and the probability density for σ is proportional to $1/\sigma$, in agreement with the previous rule. If σ is fixed, $g_{\alpha\alpha}$ is constant and the probability density for α is uniform, again in accordance with the previous rule. But if α and σ are both to be treated as unknown, the determinant is $2/\sigma^4$, and then

$$P(d\sigma\, d\alpha \mid H) \propto d\sigma\, d\alpha/\sigma^2, \tag{18}$$

instead of the previous form $d\sigma\, d\alpha/\sigma$. The effect on the t distribution would be substantial, the index being increased by $\frac{1}{2}$. Further, if the same method was applied to a joint distribution for several variables about independent true values, an extra factor $1/\sigma$ would appear for each, and the index in the t distribution for n observations would always be $\frac{1}{2}(n+1)$ whatever the number of unknowns estimated. These consequences are unacceptable. To see what (18) really says, let us suppose first that α is known to lie between fixed limits independent of σ, which corresponds fairly well to the usual practical case. Then by integration the probability density for σ is proportional to $1/\sigma^2$, and there is probability 1 that σ is less than any assigned positive value. To avoid this we should have to suppose that the range permitted to α is proportional to σ. There would be nothing self-contradictory in this, but it is not the normal state of affairs in an estimation problem.

The conclusion to be drawn is that it would be a self-consistent and generally applicable procedure to take the prior probability density proportional to $\| g_{ik} \|^{\frac{1}{2}}$ subject to the chances being differentiable functions of the parameters, and therefore that under this condition a self-consistent general theory of induction is possible. It does not, however, cover the whole ground because where it has been seen to succeed it gives the estimation prior probability, not one suitable for significance tests; and if a scale parameter σ is taken as an unknown, irrelevant to the values of the other parameters, a correcting factor is needed.

Next consider simple sampling. Denote the chance of a success by $p = \sin^2\alpha$, that of a failure by $\cos^2\alpha$. Then for variations of α

$$I_1 = (\sin\alpha - \sin\alpha')^2 + (\cos\alpha - \cos\alpha')^2 = 4\sin^2\tfrac{1}{2}(\alpha' - \alpha), \tag{19}$$

$$I_2 = (\sin^2\alpha' - \sin^2\alpha)\log\frac{\tan^2\alpha'}{\tan^2\alpha} \div 4(\alpha' - \alpha)^2. \tag{20}$$

Then the rule gives, since $0 \leqslant \alpha \leqslant \frac{1}{2}\pi$,

$$P(d\alpha \mid H) = \frac{2}{\pi}d\alpha = \frac{dp}{\pi\sqrt{\{p(1-p)\}}}. \tag{21}$$

According to the Bayes-Laplace rule the prior probability of p should be taken uniform. Haldane has suggested

$$P(dp \mid H) \propto \frac{dp}{p(1-p)}, \tag{22}$$

which has the property that the expectation of p, given the sampling ratio, is equal to the sampling ratio. The present rule is intermediate. It is easy to construct models to make any of these correspond to states called ignorance in ordinary language, but the present more precise language makes it necessary to distinguish between them.

For the comparison of two Poisson distributions of the form

$$P(m \mid rH) = e^{-r}\frac{r^m}{m!},$$

then

$$I_1 = 2-2e^{-\frac{1}{2}(\sqrt{r'}-\sqrt{r})^2}$$

leading to

$$P(dr \mid H) \propto dr/\sqrt{r}.$$

This conflicts with the rule dr/r previously adopted for the Poisson parameter, and would say that r is practically certain to exceed any suggested value. Alternatively, in most cases of the Poisson law it would be equally reasonable to take r or $1/r$ as the fundamental parameter; and the dr/r law is the only one that has the same form for both.

3. *Limitations of the rule.* The requirement that the chances shall be differentiable with respect to the parameters is not always satisfied. One case is the rectangular distribution. If the laws are that the chance of a measure in a range dx is dx/α for $0<x<\alpha$, or dx/α' for $0<x<\alpha'$, take $\alpha'>\alpha$. Then according to the first law $dP = 0$ for any interval included in $\alpha<x<\alpha'$, and I_2 is infinite. I_1 is still finite; it is, in fact,

$$I_1 = \int_0^{\alpha}\left(\frac{1}{\sqrt{\alpha'}}-\frac{1}{\sqrt{\alpha}}\right)^2 dx+\int_{\alpha}^{\alpha'}\frac{dx}{\alpha'} = 2\left(1-\sqrt{\frac{\alpha}{\alpha'}}\right).$$

For α' slightly greater than α, I_1 is of the first order in $\alpha'-\alpha$ instead of the second.

Another exceptional case is where the unknown itself is capable of taking only discrete values, as in the sampling of a finite population or in the 'tramcar' problem (Jeffreys 1939, p. 186). Here we cannot even try to differentiate. If the suggested numbers of the class are n and n', then, for a given n, I_1 is finite but takes only a discrete set of values; I_2 is infinite except for $n' = n$.

4. On inspection it is seen that most of these anomalies concern laws with a common feature; they include one or more parameters with an infinite range of possible values, and both for the normal law (with any number of parameters), the rectangular law, and the Poisson law the previous solution was quite satisfactory. These parameters are dimensional (though the Poisson parameter is itself a number, it arises as the product of two dimensional quantities, one of which is the arbitrary unit taken as a 'trial'). Suppose then that a law depends on a location parameter λ, a scale parameter σ of the same dimensions as λ, and on a number of numerical parameters α_i. Examples will be the indices in various Pearson laws and the negative binomial. It is just the occurrence of these parameters that lead in practice to a doubt about what should be taken as the best way of stating the law; for instance, whether any asymmetrical law should be stated with the first moment or the mode as location parameter. But different ways of stating the law that would ever be considered fall into the type represented by the transformation

$$\lambda' = \lambda+\sigma f(\alpha_i), \quad \sigma' = \sigma g(\alpha_i), \quad \alpha'_j = \alpha'_j(\alpha_i). \tag{23}$$

In practice, again, there is usually some more or less vague information indicating that λ and σ are practically certain to lie within wide but finite limits. With a transformation of this form the ranges permitted to λ' and $\log\sigma'/a$ (where a is any fixed quantity of the same dimensions as σ), for given α_i, will be practically the same as for λ and $\log\sigma/a$. Now

$$\frac{\partial(\lambda',\sigma',\alpha_1'\ldots\alpha_m')}{\partial(\lambda,\sigma,\alpha_1\ldots\alpha_m)} = \frac{\partial(\lambda',\sigma',\alpha_1'\ldots\alpha_m')}{\partial(\lambda,\sigma,\alpha_1'\ldots\alpha('_m)}\frac{\partial(\lambda,\sigma,\alpha_1'\ldots\alpha_m')}{\partial(\lambda,\sigma,\alpha_1\ldots\alpha_m)}$$

$$= g(\alpha_i)\left\|\frac{\partial\alpha_j'}{\partial\alpha_i}\right\|. \tag{24}$$

Hence if I_1 or I_2 is formed for changes of the α_i only, then

$$d\lambda\frac{d\sigma}{\sigma}\|g_{ik}\|^{\frac{1}{2}}\Pi d\alpha_i = d\lambda'\frac{d\sigma'}{\sigma'}\|g_{jl}'\|^{\frac{1}{2}}\Pi d\alpha_j', \tag{25}$$

in the sense that the integrals of both quantities over corresponding regions in the two cases are equal, and the same constant factor will be needed to make the total probability of all possible values equal to unity. Thus the condition of invariance for this type of transformation gives again the same type of law for the numerical parameters, but permits us to retain the previous form $d\alpha d\sigma/\sigma$ for location and scale parameters.

Most of the difficulty for laws that are not everywhere differentiable with regard to some parameter also disappears. It is well known that with some values of the indices Pearson Type I laws behave like $(x-c)^m$ near a terminus, where m may be less than 1, and that $m = 1$ is a critical value in the fitting of the law to observations. If $m > 1$, the law is differentiable at $x = c$, and the uncertainty of an estimate of c from the observations is nearly proportional to $n^{-\frac{1}{2}}$, where n is the number of observations. If $m < 1$, the law is not differentiable at $x = c$, and the uncertainty decreases faster than $n^{-\frac{1}{2}}$ (like n^{-1} for the rectangular law). The present device covers both cases, if we agree to take the termini in these laws always as $\lambda \pm \sigma$; for with fixed termini I_1 is a quadratic for small variations of the indices over the whole range of intelligibility of the laws.

Consequently the form (25) for the prior probability can be applied consistently over most estimation problems where both a scale and a location parameter have to be determined.

The outstanding exception is the case where a parameter can take only discrete values. This difficulty does not appear insuperable. Consider, for instance, the sampling of a finite population of number n, containing r members with property ϕ. This is often itself a sample of n derived from a chance x. Then

$$P(dx \mid nH) = \frac{dx}{\pi\surd\{x(1-x)\}},$$

$$P(r \mid n, x, H) = \frac{n!}{r!(n-r)!}x^r(1-x)^{n-r},$$

$$P(r\,dx \mid nH) = \frac{n!}{\pi\, r!(n-r)!}x^{r-\frac{1}{2}}(1-x)^{n-r-\frac{1}{2}}\,dx,$$

and by integration $$P(r\mid nH)=\frac{(r-\frac{1}{2})!\,(n-r-\frac{1}{2})!}{\pi r!(n-r)!},$$

which is finite both for $r=0$ and $r=n$. If used when the population is not derived from a chance x, it will still be satisfactory.

Applications to determine posterior probabilities have not been developed in the present paper, because it is known that, within any limits that anybody would ordinarily suggest, changes of the prior probability make little difference to the posterior probability in practical cases, and the modifications of existing solutions would be easy. The use of the prior probability is to provide a starting-point, and it now appears that a consistent rule for it can be stated so as to be applicable to the majority of estimation problems.

Some significance tests based on the use of I_1 and I_2 when not small have been developed, and will, it is hoped, be the subject of a further paper.

5. *Relation to χ^2 and maximum likelihood.* Suppose the possible values of the variable to be discrete; if the laws are continuous, suppose the values grouped. If there are N observations the expectations of the numbers of observations in the rth group according to the two laws are Np_r, Np'_r. Suppose that the set of parameters α_i are used as a trial hypothesis, and that the set α'_i are in fact correct. Let the observed number in the rth group be n_r. Then

$$\chi^2=\Sigma\frac{(n_r-Np_r)^2}{Np_r}, \tag{26}$$

and if N is large the expectation of χ^2, where n is the number of groups, is

$$\begin{aligned}E(\chi^2)&=n-1+\Sigma\frac{(p'_r-p_r)^2}{p_r}\\&=n-1+NI_2.\end{aligned} \tag{27}$$

Again, let α_i be taken as a set of approximate values, put $\alpha'_i=\alpha_i+\Delta\alpha_i$, and consider the estimation of $\Delta\alpha_i$ by maximum likelihood. The function to be made a maximum is

$$L=\Pi(p'), \tag{28}$$

the product being over the observed values of x. In the ordinary case the terms in $\Delta\alpha_i\Delta\alpha_k$ in $\log L$ will be

$$\tfrac{1}{2}\Delta\alpha_i\Delta\alpha_k\Sigma\frac{\partial^2}{\partial\alpha_i\partial\alpha_k}\log p'. \tag{29}$$

The expectation of the coefficient, if the correct values are α_i, will be

$$\begin{aligned}\tfrac{1}{2}N\int p\frac{\partial^2}{\partial\alpha_i\partial\alpha_k}\log p\,dx&=\tfrac{1}{2}N\int\frac{\partial^2 p}{\partial\alpha_i\partial\alpha_k}dx-\tfrac{1}{2}N\int\frac{1}{p}\frac{\partial p}{\partial\alpha_i}\frac{\partial p}{\partial\alpha_k}dx\\&=-\tfrac{1}{2}Ng_{ik},\end{aligned} \tag{30}$$

by differentiation and then integration by parts, since $\int p\,dx = 1$. Hence $-\frac{1}{2}Ng_{ik}\Delta\alpha_i\Delta\alpha_k$ is the expectation of the quadratic terms in the logarithm of the likelihood, given that the parameters have the value α'_i.

6. *Joint probabilities for several observations.* Let N observed values be derived independently from each law; their joint chances will be the products of their separate chances, and I_{1N}, I_{2N} can be defined as measures of the difference between the joint chances. Then

$$I_{1N} = \Sigma\ldots\Sigma(\sqrt{(\Pi p'_r)} - \sqrt{(\Pi p'_r)})^2. \tag{31}$$

Π refers to a product over the N values and the summations are over the values of r. Then

$$I_{1N} = 2 - 2\Sigma\Sigma(\sqrt{(\Pi p'_r)}\sqrt{(\Pi p_r)}) = 2 - 2(1 - \tfrac{1}{2}I_1)^N. \tag{32}$$

Hence $\frac{1}{N}\log(1 - \frac{1}{2}I_{1N})$ is independent of N. Also

$$I_{2N} = \Sigma\ldots\Sigma(\Pi p'_r - \Pi p)\log\frac{\Pi p'_r}{\Pi p_r} = \Sigma\ldots\Sigma(\Pi p'_r - \Pi p_r)\,\Sigma\log\frac{p'_r}{p_r}$$

$$= N\Sigma p'_r\log\frac{p'_r}{p_r} - N\Sigma p_r\log\frac{p'_r}{p_r} = NI_2 \tag{33}$$

and I_{2N}/N is independent of N.

REFERENCE

Jeffreys, H. 1939 *Theory of probability.* Oxford, Clarendon Press.

The mechanism of the hydrogen-oxygen reaction
IV. The activation energy of the initiating process

By C. F. Cullis and C. N. Hinshelwood, F.R.S.

(*Received* 28 *August* 1945)

On the basis of the theory discussed in the preceding parts of this series, a method has been devised for determining the temperature coefficient and hence the activation energy of the process responsible for starting the chains in the thermal reaction between hydrogen and oxygen. The results are consistent with the view that the production of hydrogen atoms (in bimolecular collisions between hydrogen and other molecules present) is the principal chain initiating reaction at temperatures above 560° C in vessels coated with alkali chlorides.

Introduction

The reaction mechanism suggested in Parts I and II (Willbourn & Hinshelwood 1945) as the most probable for the combination of hydrogen and oxygen above 560° C leads to the following expression for the rate of formation of water in the steady state:

$$\frac{d[H_2O]}{dt} = \frac{f_1\left\{\dfrac{1\cdot 5k_5 + 2k_6[H_2]}{k_5 + k_6[H_2]}\right\}}{1 - \dfrac{2k_2}{\Sigma k_4[M]} - \dfrac{k_6[H_2]}{k_5 + k_6[H_2]}}.$$

This may be written in the form: $2f_1 R^*$, where f_1 is the rate of chain initiation and R^* is a quantity calculable from data concerning the third explosion limit only, without reference to experimental measurements of the rates of reaction below this limit. The quantity f_1 is, of course, a function both of the concentrations of the reacting gases and of the temperature.

In Part II from the influence of the concentrations of hydrogen and oxygen on the rates of reaction at a constant temperature the probable form of f_1 was determined. This accorded best with the assumption that the dissociation of hydrogen molecules in the gas phase is the principal method of chain initiation under the conditions employed (potassium chloride-coated vessel: temperature $> 560°$ C).

The object of the present work was to determine the temperature dependence of the function f_1. Rates of reaction for mixtures containing fixed concentrations of the reacting gases have been measured at a series of different temperatures, the values of R^* for the same mixtures and at corresponding temperatures being calculated from the appropriate third explosion limit data. By division a series of values of f_1 are obtained from which an activation energy may be calculated. The results are found to be in reasonably good agreement with the conclusion that the reaction: $H_2 + M \rightarrow 2H + M$ (where M is any molecule present) is principally responsible for initiating the chains at temperatures between 560 and 600° C.

THEORETICAL CONSIDERATIONS

According to theory, assuming that the initiating mechanism consists in the gas-phase dissociation of hydrogen molecules, the rate of reaction, r (expressed in mm. of Hg per minute), is given by the expression

$$r = Bk_0[[H_2]\{Z_{H_2}[H_2]+Z_{O_2}[O_2]+Z_M[M]\}]\,R^*$$

$$= Bk_0 Z_{H_2}\left[[H_2]\left\{[H_2]+\frac{Z_{O_2}}{Z_{H_2}}[O_2]+\frac{Z_M}{Z_{H_2}}[M]\right\}\right]R^*,$$

where k_0 is the velocity constant for the initiating reaction and B is merely a factor for the conversion of units.

Now $k_0 = \text{const.}\, e^{-E_0/RT}$, where E_0 is the activation energy of the chain-initiation process; Z_{H_2} (the collision number at unit concentration for the collision of hydrogen molecules with one another) is proportional to the square root of the absolute temperature, and, since the gas mixture is of the same composition and total concentration throughout, the quantity in brackets will not vary with temperature. It is thus possible to write

$$r = \text{const.}\, e^{-E_0/RT}\sqrt{T}\,R^* \quad \text{or} \quad \ln\left(\frac{r}{R^*\sqrt{T}}\right) = \text{const.} - \frac{E_0}{RT}.$$

E_0 can thus be determined from the slope of a linear plot of the function $\ln\left(\frac{r}{R^*\sqrt{T}}\right)$ against the reciprocal of the absolute temperature.

EXPERIMENTAL METHODS

The apparatus was similar to that used by Willbourn & Hinshelwood, although a somewhat larger silica reaction vessel was employed (9 cm. long, 6·9 cm. in diameter). In the first series of experiments the walls of the vessel were thickly coated with potassium chloride; in the second series caesium chloride was used.

Calculation of the function R^*

The experimental determination of the third explosion limit at temperatures where the rate of reaction is measurable over an appreciable range of concentrations is a matter of some difficulty. The values of the third explosion limit constants have therefore been determined by experiment at a series of somewhat higher temperatures and then extrapolated; the use of a fairly large reaction vessel, however, reduced the temperature range over which this extrapolation needed to be made.

The theory of the variation of the third explosion limit with temperature has already been discussed (in Part I) by Willbourn & Hinshelwood, who carried out preliminary measurements with a gas mixture of single composition.

In the present work, complete curves connecting the third explosion-limit pressure (as defined in Part I) with the composition of the gases were obtained at different temperatures and these are shown in figures 1 (*a*)† and 1 (*b*) together with the corresponding calculated curves. In the comparatively narrow temperature range over which measurements could be made, the results were found to be in good agreement with the theory. The constant K_{24} was taken as 150 at 585° and the corresponding activation energy as 24·6 kcal. The provisional results of Willbourn & Hinshelwood for the temperature variation of the constant C are now superseded by improved values, which give $C \propto T^{1\cdot5} e^{+15,000/RT}$.

(*a*) Third explosion limit (potassium chloride-coated vessel).

A. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{150}{[\mathrm{H_2}]}}=966.$

A′. Experimental curve at 585° C.

B. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{137}{[\mathrm{H_2}]}}=1008.$

B′. Experimental curve at 580° C.

C. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{125\cdot1}{[\mathrm{H_2}]}}=1053.$

C′. Experimental curve at 575° C.

D. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{119\cdot4}{[\mathrm{H_2}]}}=1077.$

D′. Experimental curve at 572·5° C.

(*b*) Third explosion limit (caesium chloride-coated vessel).

A. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{225}{[\mathrm{H_2}]}}=1310.$

A′. Experimental curve at 606° C.

B. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{210}{[\mathrm{H_2}]}}=1353.$

B′. Experimental curve at 602° C.

C. Curve calculated according to equation: $\dfrac{[\mathrm{H_2}]+3\cdot0[\mathrm{O_2}]}{1+0\cdot40p_{\mathrm{O_2}}-\dfrac{190}{[\mathrm{H_2}]}}=1397.$

C′. Experimental curve at 598° C.

---

† The reaction vessel was recoated with potassium chloride before the measurements on the rates of reaction were made; after this had been done, it was found that the chain-breaking efficiency of the walls had decreased slightly so that the third explosion limit was depressed somewhat below the value corresponding to figure 1 (*a*). The limit curves were, however, of the same form as in the previous determination and the only constant which had altered appreciably was C. The values of the latter were found to be: 908 (585°), 949 (580°), and 991 (575°); these were used in extrapolating to lower temperatures to determine the value of the function R^*.

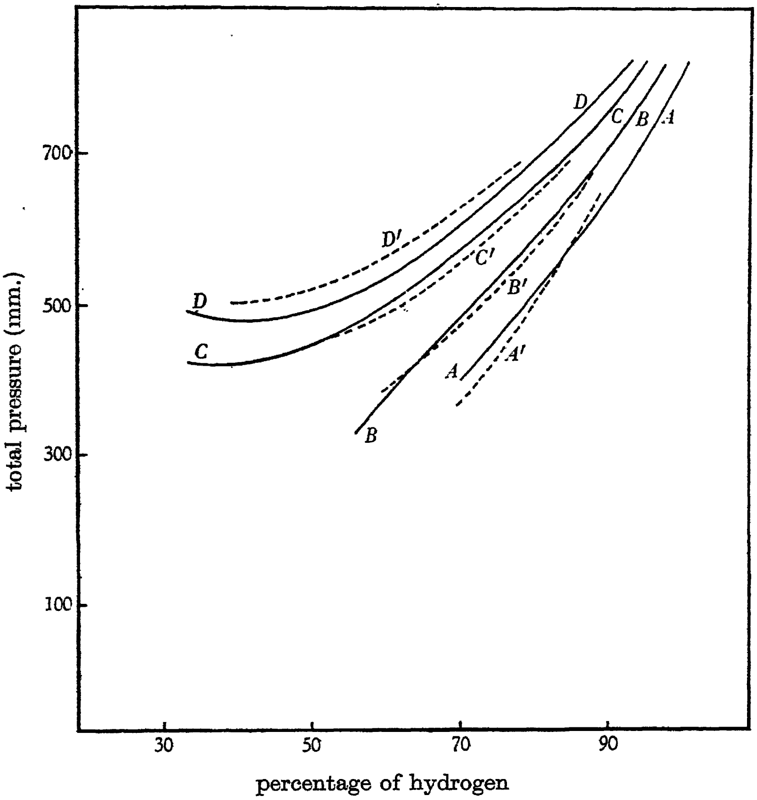


FIGURE 1*a*

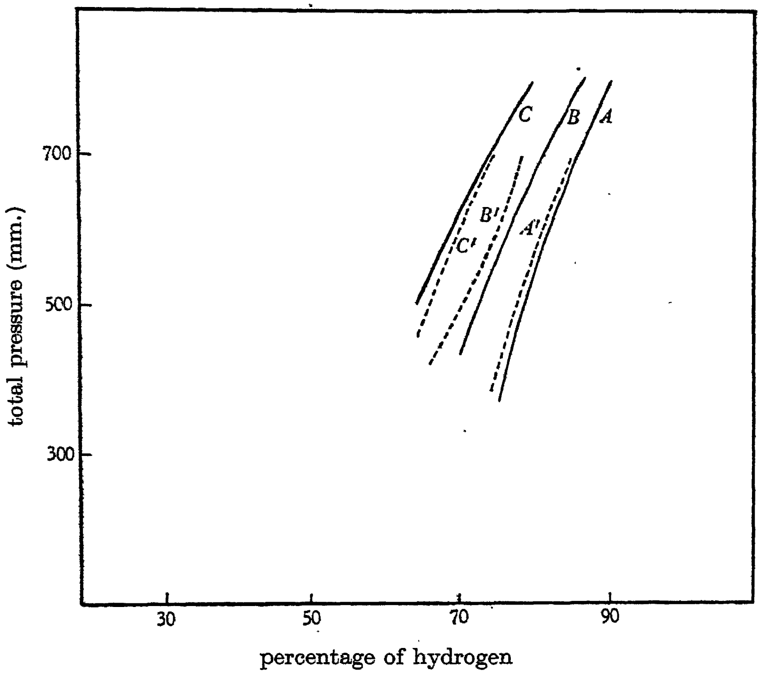


FIGURE 1*b*

The substitution of caesium chloride for potassium chloride causes a considerable elevation of the third explosion limit, on account presumably of the greater chain-breaking efficiency of the former salt; this is shown by an increase in the value of the constant C.

Owing to the ease with which the gases inflamed on mixing at the temperatures employed with caesium chloride, only relatively small portions of the third explosion limit curves could be obtained with this salt.

Measurement of the rates of reaction

The rates of reaction were measured for gas mixtures containing various definite *pressures* of hydrogen and oxygen over a range of nearly 20° C in presence of both potassium chloride and caesium chloride. In a constant-volume system, concentrations cannot be expressed as partial pressures except at a given temperature. Since R^* was to be calculated for mixtures containing fixed *concentrations* of the reacting gases, a small correction was applied to the observed rates to allow for this difference. The *corrected* rates (r') are shown in table 1 together with the corresponding calculated values of the function R^*.

It will be seen that the greater chain-breaking efficiency of caesium chloride (as compared with potassium chloride) manifests itself again here in the depressing effect on the rates of reaction.

Discussion of results

From the inset tables in figures 2 (*a*) and 2 (*b*), in which $\log\left(\frac{r'}{R^*\sqrt{T}}\right)$ is plotted against $1000/T$, it will be seen that the mean values of E_0 are approximately 110 and 92 kcal. for the results obtained with the potassium chloride- and caesium chloride-coated vessels respectively. The slopes were determined by the method of least squares. Having regard to the small temperature range over which measurements could be made, these results may be considered to be in good agreement with the conclusion that the reaction $H_2 + M \rightarrow 2H + M$ is responsible for starting the chains at the temperatures employed; the heat of reaction of the latter process (which will not differ appreciably from the activation energy) is 101 kcal.

It was on account of the great importance of the quantity E_0 that the original determination with the potassium chloride-coated vessel was followed by a similar investigation with a caesium chloride-coated vessel, which, it was hoped, would give a more accurate value for this activation energy. The latter salt appears to be even more efficient than potassium chloride in destroying HO_2 radicals. The condition that the chain-breaking efficiency of the walls approximates to unity is, therefore, more closely satisfied in presence of caesium chloride. Under these conditions, the assumption made in the derivation of the third explosion limit equation is more nearly justified. From purely theoretical considerations, therefore, the value of E_0 as determined in a caesium chloride-coated vessel should be the more reliable. In fact, however, the behaviour of the gases was somewhat erratic when the latter salt was used, and the experimental results were decidedly less reproducible than with potassium chloride. This is possibly due to the fact that caesium chloride may repossess the greater vapour pressure at the temperatures employed, which would

TABLE 1

Note. The gas compositions correspond to constant *concentrations* equivalent to the pressures recorded for the lowest temperatures of the various series.

(*a*) Calculated values of R^* and observed values of the rates of reaction in presence of potassium chloride

gas composition	temp. ° C	R^*	r' (mm.Hg/min.)
$[H_2] = 250$ mm., $[O_2] = 100$ mm.	560	8·01	1·70
	564	9·42	2·98
	567	10·88	3·87
	570	12·98	5·63
	573	16·67	9·75
	575	20·54	16·72
	578	33·55	30·35
$[H_2] = 300$ mm., $[O_2] = 100$ mm.	560	9·60	3·00
	564	11·17	5·17
	567	12·77	6·84
	570	15·00	9·88
	573	18·71	16·15
	575	22·45	26·91
$[H_2] = 350$ mm., $[O_2] = 100$ mm.	560	11·90	4·80
	564	13·93	8·16
	567	16·01	11·50
	570	18·87	15·90
	573	23·59	24·12
	575	28·32	43·20

(*b*) Calculated values of R^* and observed values of the rates of reaction in presence of caesium chloride

gas composition	temp. ° C	R^*	r' (mm.Hg/min.)
$[H_2] = 300$ mm., $[O_2] = 100$ mm.	578	5·42	2·70
	584	6·65	4·17
	589	8·55	6·42
	592	10·61	9·64
	596	15·79	17·43
$[H_2] = 350$ mm., $[O_2] = 100$ mm.	578	5·83	3·70
	584	6·91	6·16
	589	8·47	9·38
	592	10·01	13·67
	596	13·28	23·11
$[H_2] = 400$ mm., $[O_2] = 100$ mm.	578	6·48	5·10
	584	7·52	9·53
	589	8·97	13·82
	592	10·37	18·89
	596	13·16	31·34

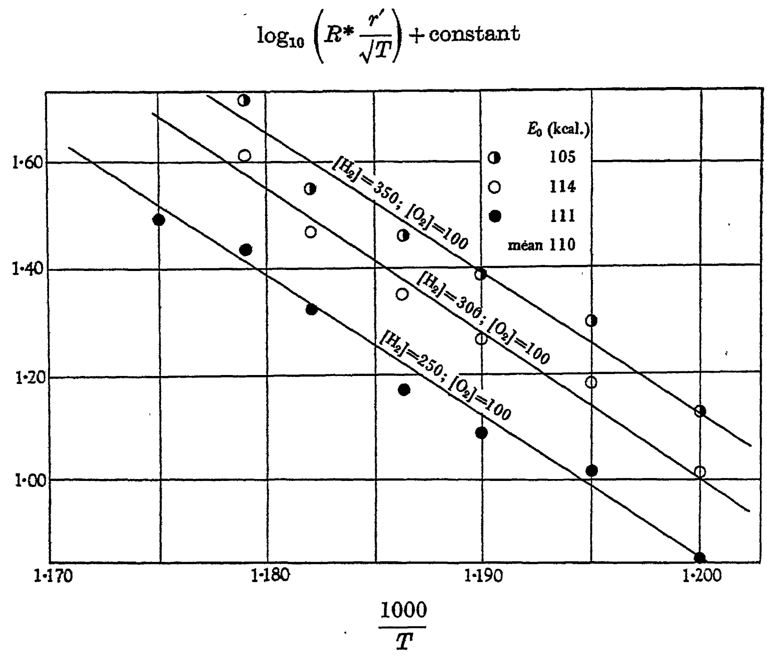


FIGURE 2*a*

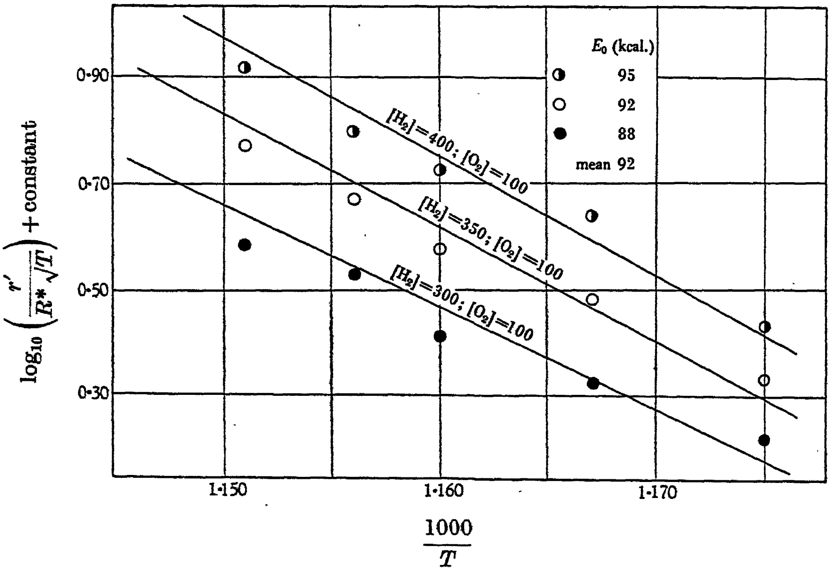


FIGURE 2*b*

increase the extent to which the salt distils over the vessel surface, altering the structure and also presumably the chain-breaking efficiency. From the experimental point of view, therefore, the value of E_0 as determined in a potassium chloride-coated vessel must be regarded as the more reliable.

In conclusion, however, it may be said that the results give general support to the conclusion already reached from the dependence of the function f_1 on the concentrations of the reacting gases, namely, that the gas phase dissociation of hydrogen molecules is the principal method of chain initiation under the conditions employed.

REFERENCE

Willbourn & Hinshelwood 1945 *Proc. Roy. Soc.* A, **185**, 353, 369.

The mechanism of the hydrogen-oxygen reaction
V. The reaction in vessels coated with alkali iodides

By C. F. Cullis and C. N. Hinshelwood, F.R.S.

(*Received* 28 *August* 1945)

Previous work on the hydrogen-oxygen combination in vessels coated with alkali halides showed that with the iodides the temperature dependence of the reaction rate is abnormal. In iodide-coated vessels the whole mechanism of the reaction is now shown to be different: the greatly reduced rate is independent of $[H_2]$, proportional to $a+b[O_2]$ and independent of added nitrogen, all in sharp contrast with what is found in chloride-coated vessels. The normal reaction is thought to be completely suppressed by minute amounts of iodine liberated into the gas phase, a residual surface reaction being measured. The chemical actions which must be assumed to occur between the iodide and the gases provide indirect evidence for the probable mode of operation of the other halide salts in controlling the hydrogen-oxygen combination.

INTRODUCTION

In Part III (Willbourn & Hinshelwood 1945) it was shown that the temperature dependence of the hydrogen-oxygen reaction between the second and third explosion limits is much smaller in vessels coated with potassium or caesium iodide than in similar vessels coated with the chlorides. It has now been shown that in iodide-coated vessels the whole nature of the reaction is radically altered. From the investigation of this abnormal reaction inferences may be drawn about the probable mode of action of the alkali halides in general.

EXPERIMENTAL RESULTS

The experimental results may be stated as follows:

(1) *The second explosion limit.* At lower temperatures the value of the limit is almost the same as that found in a chloride-coated vessel, but, as the temperature is raised, the results become increasingly erratic. They are markedly dependent upon the rate of withdrawal of the gases from the reaction vessel, the limit becoming lower as this rate is decreased. At high temperatures the lowering of the limit becomes very considerable, until finally the low-pressure explosion region is completely wiped out.

(2) *The rate of reaction between the second and third limits.* (*a*) The rate is independent of hydrogen pressure even at the highest temperatures at which measurements can be made, as shown by some typical results in table 1.

TABLE 1. DEPENDENCE OF RATE ON HYDROGEN PRESSURE IN POTASSIUM IODIDE-COATED VESSEL AT 570° C. OXYGEN PRESSURE 100 MM.

P_{H_2} (mm.)	250	300	350	400	450	500
rate (mm./min.)	6·4	6·6	6·7	6·5	6·6	6·9

(*b*) The rate increases nearly in proportion to the oxygen pressure, being represented by the expression $a + b[O_2]$, a and b being constants at a given temperature. An example is shown in table 2.

TABLE 2. DEPENDENCE OF RATE ON OXYGEN PRESSURE AT 570° C. HYDROGEN PRESSURE 300 MM.

P_{O_2} (mm.)	rate (mm./min.)	$1·6 + 0·05P_{O_2}$
100	6·6	6·6
150	9·6	9·1
200	11·6	11·6
250	13·0	14·1
300	16·7	16·6

(*c*) The rate is unaffected by additions of water vapour or of nitrogen (table 3).

TABLE 3. INFLUENCE OF NITROGEN ON RATE AT 584° C. OXYGEN 100 MM.: HYDROGEN 300 MM.

P_{N_2} (mm.)	0	50	100	200
rate (mm./min.)	13·2	13·4	14·0	13·0

(*d*) The pressure-time curves do not show the usual autocatalytic effects.

(3) *The third explosion limit.* The gases do not explode below 610° C, whereas in the same vessel coated with potassium chloride the explosion occurred at about 570° C for comparable pressures. At the high temperature in the iodide-coated bulb

the explosion seemed to be purely thermal: there was no induction period, and the rate just below the limit was very high. Added nitrogen did not lower the limiting pressures of the reacting gases.

DISCUSSION OF RESULTS

It is quite evident that the normal gas-phase reaction has been almost entirely replaced by a reaction of a quite different character. This is almost certainly due to the liberation of small quantities of iodine by the action of hydrogen on the salt and subsequent decomposition and oxidation of hydrogen iodide. Iodine acts as a very powerful inhibitor of the gas-phase reaction (Garstang & Hinshelwood 1931 *a*, *b*). This view is supported by the observation that after an explosion the condensed water was coloured faintly brown. A very minute trace of iodine would be enough to explain the effects on the second limit and the suppression of the normal reaction. On the other hand, water could continue to be formed by the cyclical formation and oxidation of hydrogen iodide (as was shown in the earlier work on the action of added iodine).

A full and detailed theory of the processes occurring in the iodide-coated vessel would be extremely complex, involving, *inter alia*, the rate of diffusion of chain carriers or iodine to or from the surface, the reversible formation of hydrogen iodide from iodine, and the oxidation of the hydrogen iodide both on the surface and in the gas phase. An approximate interpretation can, however, be indicated. Nitrogen has little effect on the rate, therefore the overall influence of diffusion on the measured rate may be neglected for the purposes of an approximate treatment, and all reactions save surface reactions left out of account. The rate is independent of $[H_2]$ and nearly proportional to $[O_2]$. This would be most simply accounted for by assuming direct attack by oxygen on a surface whose constitution is independent of $[H_2]$. If the potassium iodide surface, by the action of hydrogen, became saturated with adsorbed hydrogen iodide, then the required condition would be fulfilled. The formation of potassium hydride and hydrogen iodide in the surface layer is quite probable, and the frequency of occurrence of the latter in the mosaic of atoms, ions and adsorbed molecules could be nearly independent of hydrogen. The oxidation of the hydrogen iodide would lead to the escape of iodine into the gas phase where its extraordinarily powerful inhibiting effect would be manifested.

Although it would be scarcely profitable to work out the full details of the mechanism, the general result is of interest in that it supports the view that all salts exert their effects essentially by their ability to enter into definite chemical reactions of the following type (cf. Part III):

$$KX + H = K + HX, \quad KX + H = KH + X,$$

X being a halogen atom.

Caesium iodide would obviously be expected to behave like potassium iodide.

With chlorides, hydrogen chloride would be formed though to an extent which, in view of the considerations mentioned in Part III, would be difficult to predict.

It would be less susceptible to oxidation and decomposition, and even if appreciable amounts of chlorine escaped into the gas phase, their inhibiting effect would be very much less marked (as has been directly shown), since chlorine atoms continue the well-known Nernst chain. This means that chlorine is not an effective remover of hydrogen atoms. Thus we can have the inhibiting effect of the salt surface without the accompanying effects of powerful chain breakers in the gas phase.

REFERENCES

Garstang, W. L. & Hinshelwood, C. N. 1931 *a* *Proc. Roy. Soc.* A, **130**, 640.
Garstang, W. L. & Hinshelwood, C. N. 1931 *b* *Z. phys. Chem.* Bodenstein-Festband, p. 656.
Willbourn, A. H. & Hinshelwood, C. N. 1945 *Proc. Roy. Soc.* A, **185**, 376.

The drag of cylinders in fluids at slow speeds

BY C. M. WHITE, *Imperial College, London*

(*Communicated by R. Bagnold, F.R.S.—Received* 7 *April* 1945)

Experiments with wires falling sideways in viscous fluids are described. At very slow speeds the resistance is found to be independent of the density of the fluid when the fluid is finite in extent. In such purely viscous motion the boundary distance controls the pattern and Reynolds's number is without significance. At higher speeds the resistance becomes independent of this distance and agrees with Lamb's solution for a cylinder in an infinite fluid. At still higher speeds the experiments close the gap between the upper limit of validity of Lamb's approximations and the lower end of the wind-tunnel experimental range. The present experiments include Reynolds's numbers ranging from 10^{-5} to 10^{+2}: at low speed the influence of distant boundaries is unexpectedly great, and Lamb's formula is much more restricted in scope than usually thought; when the Reynolds number is 10^{-4} the presence of boundaries 500 diameters away multiplies the drag two-fold.

Notation

P = drag per unit length of cylinder (MT^{-2}).
τ = P/d = net weight of cylinder$/dL$ ($ML^{-1}T^{-2}$).
L = length of cylinder (L).
d = diameter of cylinder (L).
D = diameter of container (L).
v = speed (LT^{-1}).
μ = viscosity of fluid ($ML^{-1}T^{-1}$).
ρ = density of fluid (ML^{-3}).
R_1 = Reynolds number, $\rho vd/\mu$.
R_2 = Reynolds number, $\rho vD/\mu$.

Some years ago I noticed that Lamb (1) showed analytically that the drag of a cylinder moving laterally in a viscous fluid was

$$\frac{\tau d}{\mu v}=\frac{5{\cdot}46}{\log 7{\cdot}4\mu/\rho v d},\tag{1}$$

which seemed to conflict with other examples of fluid motion which all conformed with a relation of the type

$$\frac{\tau d}{\mu v}=\text{constant}\tag{2}$$

when the Reynolds number was small enough. At zero speed both (1) and (2) are of course indistinguishable, but with increasing speed Lamb's result begins to deviate from (2) much sooner than would be expected by comparison with other flow problems.

Accordingly, I experimented with metal wires falling in gylcerine solutions, and found that, except at higher speeds, Lamb's formula applies only in an infinite or nearly infinite fluid, and that in any practical case when the fluid is limited by outer boundaries there is a very marked deviation from (1). Most of the experimental results conformed closely with (2), the value of the constant depending on the distance of the outer boundaries. Lamb's formula seems exact enough numerically, but at the slower speeds it applies only when the boundaries are very far away indeed; for example, when $R = 0{\cdot}001$, boundaries 500 diameters away completely determine the drag, and it is not until they are some 10,000 diameters away that their influence disappears.

The first experiments were arranged to see whether (2) ever applied, that is, to see whether there was a range of speeds in which the fluid inertia could be neglected. According to (1) this is never so, even at the lowest speeds. Eleven wires of various diameters were allowed to fall in two different glycerine solutions, the second 2·2 times as viscous as the first. In this way two terminal speeds for each wire were determined under conditions giving Reynolds's numbers approximately in the ratio of 4·7. The same containing jar was used, so maintaining strict geometrical similarity in the form of the boundaries, and as the density of the solutions was almost identical each wire experienced the same drag in both solutions. The results are given in table 1. For each wire there are only two variables, viscosity and speed; if these are inversely proportional then the coefficient $\tau d/\mu v$ must be strictly independent of R_1. The last column but one in table 1 provides good evidence that this is so, the ratios of the speeds being almost exactly in inverse ratio with the viscosities. In this column, unity represents complete independence of R_1: excluding one doubtful value, the remaining ten wires average 1·008, which, over the range of Reynolds's numbers from 0·13 to 0·0001, corresponds to a total variation of less than 4 % in the resistant coefficient despite the 1300 to 1 change in R_1. According to (1) the drag should have changed 2·7 times. So the experiment clearly proves that at these speeds Lamb's result is inapplicable even when the fluid extends some 150 diameters from the wire. Further, it demonstrates the existence of a purely viscous regime in accordance with (2).

TABLE 1. COMPARISON OF TERMINAL SPEEDS OF COPPER WIRES FALLING IN TWO GLYCERINE SOLUTIONS OF DIFFERENT VISCOSITIES

Solution A, $\mu_A = 6{\cdot}0$ g.cm.$^{-1}$sec.$^{-1}$; solution B, $\mu_B = 13$ g.cm.$^{-1}$sec.$^{-1}$; density, both A and B approx. 1·26 g.cm.^{3}. Containing jar, 5·86 cm. diam., 45 cm. deep.

nominal diam. of wire mm.	length of wire cm.	Reynolds number $1000\frac{\rho vd}{\mu}$		$\frac{\text{speed A}}{\text{speed B}}$	$\frac{\text{viscosity B}}{\text{viscosity A}}$	$\frac{v_B\mu_A}{v_A\mu_B}$	ditto calculated from (1)
		solution A	solution B				
1·44	3	127·5	28·2	2·08	2·17	1·04	1·37
0·818	3	30·7	6·5	2·18	2·17	0·99$_5$	
0·818	1	27·1	5·3	2·24	2·26	1·01	
0·567	3	8·8	1·82	2·23	2·17	0·97$_3$	
0·567	1	7·4	1·66	2·23	2·26	1·01$_3$	
0·379	3	3·5	0·78	2·06	2·17	1·05	1·20
0·379	1	3·4	0·68	2·21	2·24	1·01	
0·286	3	1·63	0·345	2·18	2·16	0·99$_4$	
0·286	1	1·54	0·319	2·17	2·22	1·02$_4$	
0·190	3	0·48?	0·125	1·99?	2·15	(1·08?)	
0·190	1	0·50	0·105	2·21	2·15	0·97$_2$	1·16
					av.	1·008	

An earlier attempt to do this failed, and gave the false impression that (1) was valid, for instead of varying the Reynolds number by varying the viscosity of the fluid, the more obvious way, by using different sizes of wire in the same fluid, was tried. A wider range of Reynolds's number was certainly obtained, but the results gave quite a false impression. The change in geometrical proportions (wire diameter/jar diameter, wire length/wire diameter) led to changes in the drag coefficient such that the Reynolds number seemed to be having just the effect predicted by (1). Luckily a fortuitous change in viscosity drew attention to the error.

The existence of such a regime was also confirmed visually by observing the motion. This was done by first dropping a number of heated or dye-covered balls in the glycerine. These left vertical hair-like lines behind them. Then a wire was dropped, and the distortion of the field about the slowly moving wire could be seen. The flow pattern was quite symmetrical above and below the falling wire, and at a rapid glance the wire might just as well have been moving upwards. Such reversibility is only possible when the motion is controlled by viscosity alone or by density alone, since the one involves time and the other does not.

With inertia thus eliminated the only independent variables which can influence the value of the resistance coefficient $\tau d/\mu v$ are the ratios L/d and d/D. The former is conveniently regarded as partly a local end-effect to be eliminated by a correction. Its value was found by comparing the terminal speeds of wires of different lengths. The results in table 2 show that as the wire is shortened its terminal speed at first rises and then falls again. The quantity L/v in the fifth column represents the total drag of each wire reduced to constant speed. This is plotted in figure 1 against L, and the curve extrapolated to $L = 0$, where L/v is seen to be 0·43 sec., which presumably represents the drag at unit speed when the whole of the drag is end-effect. This is a few per cent greater than the drag of a sphere of the same diameter as the wire, a convenient form in which to express the end-correction. The corresponding correction is approximately

$$10\frac{d}{L}\left(1+2{\cdot}4\frac{d}{D-L}\right), \tag{3}$$

to be subtracted from the observed value of $\tau d/\mu v$. This was used, but only when Reynolds's number was less than unity: when $R > 5$ an additive correction was used increasing the drag in the ratio $(1+4{\cdot}8d/L)$.

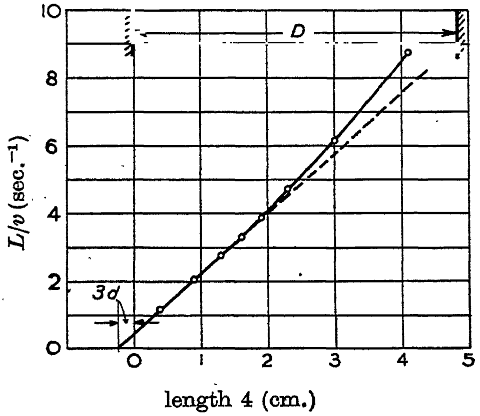


FIGURE 1. Influence of length of wire, $d = 0{\cdot}818$ mm.

TABLE 2. EFFECT OF LENGTH OF WIRE

Diameter of wire, 0·818 mm.; mass, 0·0463 g./cm.; diameter of jar, 4·8 cm.; viscosity, 19·0 g.cm.$^{-1}$sec.$^{-1}$; density, 1·26 g.cm.$^{-3}$ approx.

					drag coefficient $\tau d/\mu v$	
length L cm.	L/d	L/D	speed v cm./sec.	L/v	observed	corrected
4·1	50·1	0·85	0·470	8·73	4·37	4·11
3·0	36·7	0·62	0·486	6·17	4·22	3·91
2·3	28	0·48	0·488	4·71	4·21	3·81
1·9	23	0·38	0·493	3·85	4·17	3·69
1·6	19·6	0·33	0·486	3·30	4·22	3·66
1·3	16	0·27	0·477	2·72	4·31	3·63
0·9	11	0·19	0·447	2·02	4·60	3·61
0·4	4·9	0·08	0·345	1·16	5·95	3·73

The main experiments were to determine $\tau d/\mu v$ as a function of d/D, and incidentally to define the limits within which purely viscous motion occurred. The experimental method was much the same as before, though the range of the variables was extended to include the following table.

Range of main experiments

fluid	viscosity g.cm.$^{-1}$sec.$^{-1}$	density g.cm.$^{-3}$
glycerine and water	5·7–19	1·26
‘Arctic’ lubricating oil and paraffin	0·6–1·0	0·9
water	0·011	1·0

	least	greatest	ratio
wire diameter	0·121 mm.	1·442 mm.	12
wire, mass per metre	0·1022 g.m.$^{-1}$	14·41 g.m.$^{-1}$	140
container diameter	1·22 cm.	7·2 cm.	5·9
container diameter/wire diameter	8·5	600	70
terminal speeds	0·018 cm.sec.$^{-1}$	30 cm.sec.$^{-1}$	1700
observed distance of fall	1 cm.	1 m.	—
Reynolds's number	$1\cdot 4 \times 10^{-5}$	2×10^{2}	14,000,000

The results for those speeds with purely viscous motion are given in figure 2, in which the drag coefficient $\tau d/\mu v$, corrected for end-effects, is plotted against the boundary nearness ratio d/D. Despite a rather noticeable scatter of the points, and this nowhere exceeds the known experimental uncertainty, there is clearly seen an upward trend from left to right, the resistance coefficient increasing threefold as d/D increases from 0·0017 to 0·12. A logarithmic type of curve can represent such a relation, and with the origin arbitrarily taken at $d/D = 0\cdot 8$ (infinite drag when wire fouls container) the most reliable of the test points give

$$\frac{\tau d}{\mu v} = \frac{6\cdot 3}{\log 0\cdot 8D/d}, \qquad (4)$$

which is shown in figure 2 in satisfactory agreement with most of the remaining points. The last two points to the right are unreliable as the end-correction was excessive, being 20 and 30 % respectively.

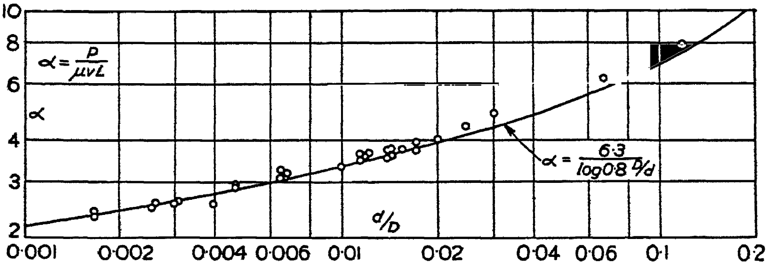


Figure 2. The drag of cylinders in purely viscous range. Only those tests in which $\frac{\rho v d}{\mu} < 3$ are plotted here.

When the Reynolds number becomes great, inertia is no longer negligible and the drag begins to be influenced by density, so (4) must then fail. For each particular Reynolds number, failure occurs when the boundaries are sufficiently far away, since with distant boundaries the drag according to (4) becomes smaller than that in an infinite fluid, taking into account, as Lamb does, the inertia terms. Thus at every speed as the boundaries recede, there must be some distance at which (4) gives place to Lamb's formula:

$$\frac{\tau d}{\mu v} = \frac{5{\cdot}46}{\log 7{\cdot}4\mu/\rho v d}. \tag{1}$$

The transition from (4) to (1) is very gradual, and is best seen when the test points for particular values of d/D are plotted against R_1 as in figure 3. In the experiments more than twenty different values of d/D were used, but to avoid confusion the results are plotted in figure 3 under only five values of d/D, any necessary small adjustments being made on the basis of (4). Points previously given in figure 2 are included together with those at higher speeds. In accordance with (4) separate horizontal lines are obtained for each value of d/D in the region A; here the drag is in no way influenced by the density of the fluid. As the speed increases to the right, inertia begins to shield the wires from the distant boundaries, and the drag is the same as that in an infinite fluid: curve B, i.e. equation (1) (Lamb) then applies. Towards the right at $R_1 = 0{\cdot}7$, Lamb's curve begins to rise too steeply, so this is the upper limit of the validity of his approximations. To the extreme right, the points tend towards the curve D which represents the lower end of the wind-tunnel experiments of Relf (1914). There is some discrepancy here, the drags even after correcting for end-effects being 5–15 % less than Relf obtained, but at still higher speeds beyond the diagram the agreement was good.

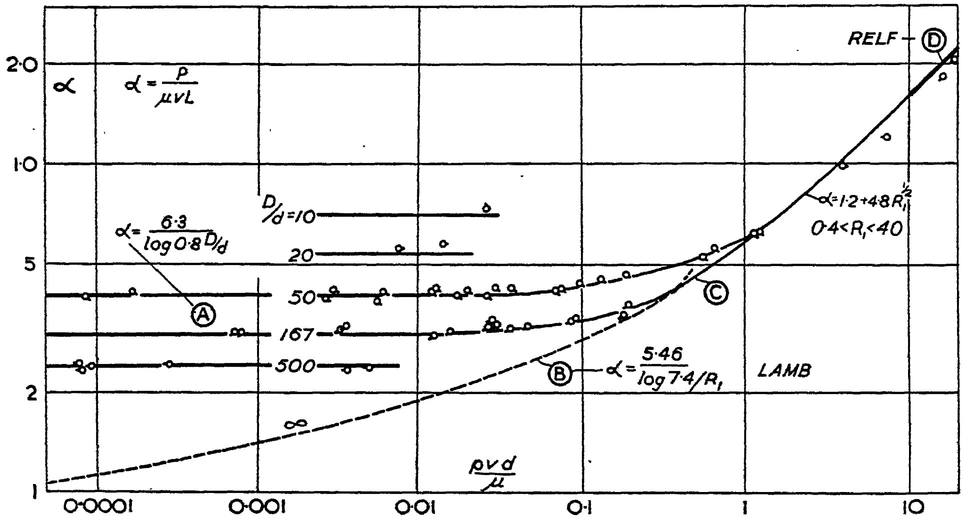


FIGURE 3. The drag of cylinders in the transition range between purely viscous motion and motion involving both viscosity and density corrected for end effects.

Most of the data of figures 2 and 3 can be expressed as a single curve after slightly changing the origin of (4). The latter is purely empirical, and the two numerical constants in it can be varied somewhat without spoiling the fit. With the origin at $0{\cdot}43D/d$, the viscous formula (4) becomes

$$\frac{\tau d}{\mu v} = \frac{5{\cdot}46}{\log 0{\cdot}43D/d}, \tag{4a}$$

which may be written as $$\frac{\mu v}{\tau d} - 0{\cdot}183 \log \frac{D}{d} = -0{\cdot}067, \tag{4b}$$

and Lamb's equation (1) can also be put in this form, thus

$$\frac{\mu v}{\tau d} - 0{\cdot}183 \log \frac{D}{d} = -0{\cdot}183 \log \frac{\rho v D}{7{\cdot}4\mu}, \tag{1a}$$

so disclosing both (4*b*) and (1*a*) as particular cases of

$$\frac{\mu v}{\tau d} - 0{\cdot}183 \log \frac{D}{d} = \phi\left(\frac{\rho v D}{\mu}\right).$$

This has been plotted in figure 4, where the points are seen to conform well enough to a single curve. The distant boundaries are seen to be the controlling factor when $\dfrac{\rho v D}{\mu} < 3$, whereas they are of no account and Lamb's (1) applies when $\dfrac{\rho v D}{\mu} > 40$.

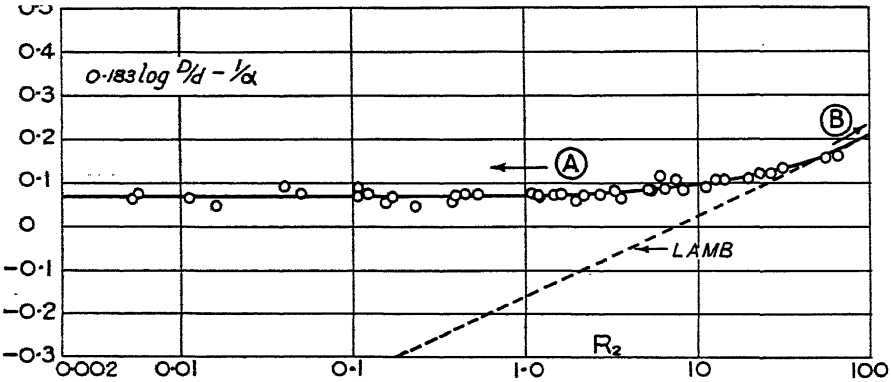


FIGURE 4. The resistance of cylinders as function of R_2; an approximate correlation of the viscous and semi-viscous regimes *A* and *B* of figure 3.

Upper limit to *A* - - - $R_2 = 3$. $R_2 = \rho v D/\mu$.
Lower limit to *B* - - - $R_2 = 40$. $R_1 = \rho v d/\mu$.
Upper limit to *B* - - - $R_1 = 0{\cdot}7$ $\alpha = \rho/\mu v L$.

So for cylinders moving at the same Reynolds number $\rho v d/\mu$, two distinct resistance laws have been found; and this may be of interest in connexion with hot-wire anemometry, particularly when speeds near a surface are to be measured. The motion, and presumably the calibration, of a hot wire will change when its distance is less than $20\mu/\rho v$ from the wall.

As one would expect, the shape of the distant walls influences the drag. A cylinder midway between plane boundaries, distant *H* from each other, falls a little faster than in a circular containing jar whose diameter equals H. The difference is no longer detectable when H is made equal to $0\cdot 8D$. Various wires and ebonite rods falling in medicinal paraffin between plates mostly 15 cm. wide and 1 or $2\frac{1}{2}$ cm. apart gave the following:

TABLE 3. CYLINDERS FALLING BETWEEN PARALLEL PLATES

diam. mm.	H/d	R_1	R_2	L/d	α observed	α corrected for ends	$\alpha \log H/d$
0·192	132	0·004	0·52	520	3·20	3·17	6·75
0·280	91	0·010	0·91	107	3·12	3·03	5·94
0·192	50	0·003	0·15	160	3·67	3·61	6·13*
0·540	47	0·07	3·24	93	4·22	4·11	6·87
0·280	34	0·008	0·26	125	4·06	3·98	6·10
0·540	17·6	0·05	0·91	65	5·44	5·21	6·48
4·32	5·9	0·30	1·77	19	10·8	9·38	7·2
4·32	2·2	0·12	0·26	10·3	27·1	21·0	7·15

* α increased 9 % as L/d increased from 53 to 317.

In the last column $\alpha \log H/d$ varies from 5·9 to 7·2, due mainly to experimental uncertainties: excluding the two highest values, the mean is 6·4. So, for motion between parallel plates

$$\frac{P}{\mu v L} = \frac{6\cdot 4}{\log H/d}. \qquad (4c)$$

Bairstow, Cave & Lang (1922), by successive approximation, obtained a stream function representing upward flow past a stationary cylinder midway between stationary plates five diameters apart. They found $\alpha = 0\cdot 71$. They remark that some of the terms in the stream function are much the same when the cylinder moves and the fluid and walls are at rest. However, the main terms are different: in their case the velocity distribution far from the cylinder is parabolic, which makes the maximum speed $\frac{3}{2}$ times the mean, and the speed of the cylinder relative to the mean fluid is two-thirds of its speed relative to the fluid straight ahead: allowing for this, the 7·1 becomes 10·6 when, as in the present experiments, v denotes speed of cylinder relative to mean fluid. Also when the stress on the outer plates is removed the pattern near the cylinder relaxes somewhat and the drag diminishes perhaps by 16 %, giving $\alpha = 9\cdot 0$ finally. Formula (4c), when extrapolated to $H/d = 5$, also gives $\alpha = 9\cdot 0$, whereas the two nearest experiments when interpolated to $H/d = 5$ give 10·2, but these were difficult ones and may be 20 % in error.

REFERENCES

Bairstow, Cave & Lang 1922 *Proc. Roy. Soc.* A, **100**, 394.
Lamb, H. 1930 *Hydrodynamics*, 5th ed. art 343, p. 583.
Relf, E. F. 1914 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 102.

The theory of axisymmetric turbulence

By G. K. Batchelor, M.Sc.(Melb.)*

Division of Aeronautics, Council for Scientific and Industrial Research, Australia

(*Communicated by Sir Geoffrey Taylor, F.R.S.—Received* 22 *August* 1945)

This paper discusses a type of turbulence in a uniform stream which is next to isotropic turbulence in order of simplicity. Instead of spherical symmetry, or isotropy, axially symmetical turbulence possesses symmetry about an axis which in practice is usually the direction of mean flow. The analysis is developed with the aid of invariant theory, as suggested by a previous paper by Robertson. The form of the fundamental velocity correlation is obtained, and scales of axisymmetric turbulence are defined.

The results of greatest practical interest concern the time rates of change of the mean squares of the lateral and longitudinal velocity components. The rates of change involve two terms, the first representing viscous dissipation, and the second representing a transfer of energy from one component to the other due to the finite correlation between the velocity and pressure at neighbouring points. The effect of the velocity-pressure correlation is to bring the two velocity components towards equality, while the effect of the viscous dissipation will only be towards equality if an inequality between the curvatures at the origin of two particular velocity correlation coefficient curves, both of which are measurable, is obeyed. The rates of change of the mean squares of the vorticity components are also obtained.

1. AXISYMMETRIC TURBULENCE

The theory of turbulence in a uniform stream has been developed, with some success, for the particular case of *isotropic* turbulence. The notion of isotropy was introduced by Taylor in 1935, and is defined by the condition that the temporal mean value of any function of the velocity components of the turbulence and their derivatives, defined in relation to a particular set of axes, is unaltered if the axes are rotated or reflected in any manner. Since the stream is uniform it is possible to assume that the turbulence is *homogeneous*, i.e. independent of position in the fluid. The turbulent fluctuations are assumed to be so rapid that no significant change in mean values occurs over the time required for averaging.

The assumption of isotropy greatly simplifies the discussion of turbulence and many results have been obtained, perhaps the most important of which relates to the rate of decay of the mean square of the velocity fluctuation. However, from the experimental point of view, isotropic turbulence is not naturally occurring, nor is it particularly easy to produce. The turbulence produced in the wake of a mesh of regularly spaced rods or wires is known to be isotropic, but, on the other hand, the turbulence passing *through* such a mesh is not isotropic (Batchelor 1945), and experiments on isotropic turbulence must therefore be made under conditions for which the former turbulence is predominant. The stream of air produced in a wind tunnel is nearly always non-isotropic, the velocity fluctuations parallel and perpendicular to the stream direction being unequal.

---

* Now at Trinity College, Cambridge.

A type of turbulence which is next to isotropic in order of simplicity, but which is more commonly occurring in practice, is *axisymmetric* turbulence (in contradistinction to the spherical symmetry of isotropic turbulence). The mean value of any function of the velocities and their derivatives must in this case be independent of arbitrary rotations of the axes about a given direction (which in all practical realizations will be the direction of mean flow) and of reflexions in planes through that direction. In this paper axisymmetric, homogeneous turbulence is examined in order to derive properties analogous to those found for isotropic turbulence. This proves to be partially possible, despite the larger number of arbitrary functions which enters the theory.

The analysis uses as extensively as possible the results of invariant theory; this was prompted by the elegance of H. P. Robertson's paper (1940) employing invariant theory in the examination of isotropic turbulence.

2. AXISYMMETRIC TENSORS AND FORMS

It was shown by de Karman & Howarth (1938) that the various double correlations between the components of velocity fluctuations can be considered as the components of a tensor. Tensors of different orders enter the theory at many points so that it will be convenient here to consider a typical tensor $F_{ij\dots}$ of arbitrary order which is the temporal mean of the product of simultaneous velocity components and/or their derivatives at no more than two points in the fluid.* Suppose for the moment that $F_{ij\dots}$ is the mean of the product of velocity components $u_i, u_j, \dots$, some of which refer to the point P of co-ordinates x_i and others to P' of co-ordinates x'_i. This tensor $F_{ij\dots}$ is relevant to an inquiry about $F(\mathbf{a}, \mathbf{b}, \dots)$ which is the mean product of velocity components $u_i a_i, u_j b_j, \dots$ taken in the directions given by the arbitrary unit vectors $\mathbf{a}, \mathbf{b}, \dots$. Hence

$$\begin{aligned} F(\mathbf{a}, \mathbf{b}, \dots) &= \overline{u_i a_i u_j b_j \dots} \\ &= \overline{u_i u_j \dots} a_i b_j \dots \\ &= F_{ij\dots} a_i b_j \dots \qquad (2\cdot1) \end{aligned}$$

(repeated indices indicating summation, and the over-bar indicating a mean).

Now in the case of axisymmetric turbulence, $F(\mathbf{a}, \mathbf{b}, \dots)$ should be unchanged by an arbitrary rotation of the axes of reference about a given direction, which is taken to be defined by the unit vector $\boldsymbol{\lambda}$, or by reflexions of the axes in planes through $\boldsymbol{\lambda}$. It will be convenient to remove the restriction on the type of rotation or reflexion, and it is therefore decided, as a *definition*, that for axisymmetric turbulence, $F(\boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \dots)$ is unchanged by arbitrary rotations or reflexions of the axes of reference.* $F(\boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \ldots)$ is then said to be an axisymmetric form and the associated tensor $F_{ij\ldots}$ is an axisymmetric tensor. The relation between isotropic and axisymmetric turbulence is thus clear; the isotropic form F must remain unaltered by arbitrary rotations or reflexions of the configuration formed by the vectors $\mathbf{a}, \mathbf{b}, \ldots$ and $\boldsymbol{\xi}$, where $\xi_i = x'_i - x_i$, while for the corresponding axisymmetric form the vector $\boldsymbol{\lambda}$ is added to the configuration which must remain unchanged. Figure 1 shows the configuration for the case of only two arbitrary vectors, **a** at P, and **b** at P'.

---

* Scalar terms, such as the pressure p, or $\dfrac{\partial u_i}{\partial x_k}\dfrac{\partial u_k}{\partial x_i}$, may also occur in the product. Such terms contribute nothing to the tensorial character of the product, but it will be noticed that the presence of a scalar term renders it possible for all the velocities and velocity derivatives to refer to the same point.

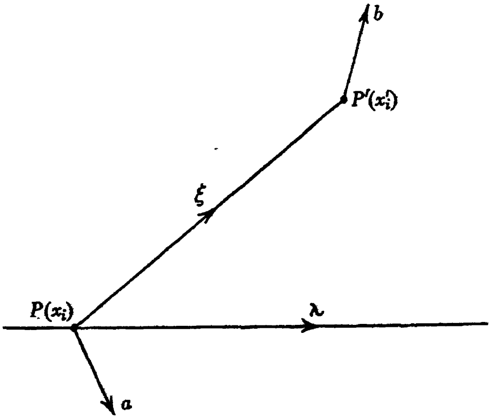


FIGURE 1

The results of invariant theory can to now be used obtain information about the dependence of axisymmetric tensors on the vectors $\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \ldots$. The general problem is: to determine the function $F(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \ldots)$ of the form (2·1) which is an invariant under an arbitrary rotation or reflexion (i.e. under the full rotation group) of the vector configuration formed by $\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \ldots$. As pointed out by Robertson (1940), it is a known result of invariant theory that such a function must depend only on the fundamental invariants satisfying the same condition, viz. on the scalar products $(\boldsymbol{\xi}.\boldsymbol{\xi}), (\boldsymbol{\xi}.\boldsymbol{\lambda}), \ldots$ formed from any two of the vectors $\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \ldots$. Choosing the combinations of scalar products which conform to the relation (2·1) thus gives the most general form of $F(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}, \ldots)$.

To proceed further one must particularize by limiting the order of the tensors. The following results are obtained for first- and second-order tensors.

Linear form

The relation (2·1) becomes in this case

$$L(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}) = L_i a_i, \tag{2·2}$$

and the scalar products available for the construction of the axisymmetric form $L(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a})$ are $(\boldsymbol{\xi}.\boldsymbol{\xi})$, $(\boldsymbol{\lambda}.\boldsymbol{\lambda})$, $(\mathbf{a}.\mathbf{a})$, $(\boldsymbol{\xi}.\boldsymbol{\lambda})$, $(\boldsymbol{\lambda}.\mathbf{a})$ and $(\mathbf{a}.\boldsymbol{\xi})$. Hence

$$L(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}) = M(\boldsymbol{\xi}.\mathbf{a}) + N(\boldsymbol{\lambda}.\mathbf{a}), \tag{2·3}$$

---

* This definition is not quite identical with that suggested earlier, since it now includes reflexion in planes normal to $\boldsymbol{\lambda}$.

where M and N are arbitrary functions of $(\boldsymbol{\xi}.\boldsymbol{\xi})$, $= r^2$, and $(\boldsymbol{\xi}.\boldsymbol{\lambda})$, $= \mu r$ say, and

$$L_i = M\xi_i + N\lambda_i. \tag{2·4}$$

It will also be useful for future reference to record the divergence of L_i. Noting that

$$\begin{aligned}\frac{\partial}{\partial \xi_i} &= \frac{\partial r}{\partial \xi_i}\frac{\partial}{\partial r} + \frac{\partial \mu}{\partial \xi_i}\frac{\partial}{\partial \mu} \\ &= \frac{\xi_i}{r}\frac{\partial}{\partial r} + \left(\frac{\lambda_i}{r} - \frac{\mu \xi_i}{r^2}\right)\frac{\partial}{\partial \mu},\end{aligned} \tag{2·5}$$

the divergence of L_i is found from (2·4) to be the scalar

$$\frac{\partial L_i}{\partial \xi_i} = 3M + r\frac{\partial M}{\partial r} + \mu\frac{\partial N}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial N}{\partial \mu}. \tag{2·6}$$

In isotropic turbulence, for which N is zero, the property of zero divergence can thus only be consistent with regularity of L_i if M, and consequently L_i, is identically zero (de Karman & Howarth 1938). However, this is not true of axisymmetric first-order tensors.

Bilinear form

Following Robertson's notation again, (2·1) becomes

$$Q(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}) = Q_{ij}a_i b_j, \tag{2·7}$$

and consideration of the available scalar products shows

$$Q(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}) = A(\boldsymbol{\xi}.\mathbf{a})(\boldsymbol{\xi}.\mathbf{b}) + B(\mathbf{a}.\mathbf{b}) + C(\boldsymbol{\lambda}.\mathbf{a})(\boldsymbol{\lambda}.\mathbf{b}) + D(\boldsymbol{\lambda}.\mathbf{a})(\boldsymbol{\xi}.\mathbf{b}) + E(\boldsymbol{\xi}.\mathbf{a})(\boldsymbol{\lambda}.\mathbf{b}), \tag{2·8}$$

where A, B, C, D and E are arbitrary functions of r^2 and μr. The associated second-order axisymmetric tensor is

$$Q_{ij} = A\xi_i\xi_j + B\delta_{ij} + C\lambda_i\lambda_j + D\lambda_i\xi_j + E\lambda_j\xi_i, \tag{2·9}$$

where
$$\begin{aligned}\delta_{ij} &= 0 \quad \text{if} \quad i \neq j, \\ &= 1 \quad \text{if} \quad i = j.\end{aligned}$$

It is apparent from figure 1 that

$$Q(\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{a}, \mathbf{b}) = Q(-\boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{b}, \mathbf{a}), \tag{2·10}$$

true, in fact, for any two-point correlation between two vector quantities of the same kind, with no scalars occurring. In the case of axisymmetric forms and tensors it leads to the conclusion that A, B and C depend only on r^2 and $\mu^2 r^2$, and that

$$D(-r) = -E(r), \quad E(-r) = -D(r), \tag{2·11}$$

or, alternatively,
$$D(-\mu) = -E(\mu), \quad E(-\mu) = -D(\mu). \tag{2·12}$$

Hence, if a member of this class of correlation tensors is symmetrical in its indices, D and E are identical and are odd functions of μr. (An isotropic tensor, for which C, D and E are zero, is *ipso facto* symmetrical.)

The divergence of Q_{ij} will be a first-order axisymmetric tensor and therefore expressible in the form (2·4). From (2·9) it is found that

$$\frac{\partial Q_{ij}}{\partial \xi_j} = \xi_i\left(4A + r\frac{\partial A}{\partial r} + \frac{1}{r}\frac{\partial B}{\partial r} - \frac{\mu}{r^2}\frac{\partial B}{\partial \mu} + \mu\frac{\partial E}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial E}{\partial \mu}\right)$$
$$+\lambda_i\left(\frac{1}{r}\frac{\partial B}{\partial \mu} + \mu\frac{\partial C}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial C}{\partial \mu} + 3D + r\frac{\partial D}{\partial r} + E\right). \quad (2{\cdot}13)$$

If Q_{ij} is a solenoidal tensor, i.e. has vanishing divergence with respect to, say, the index j, this expression is zero. Since ξ_i can take infinitely many values for constant values of r, μ and λ_i, both terms in square brackets vanish. Since A, B and C are even functions of μr and r for correlations not involving scalars, it then follows that the expressions

$$\mu\frac{\partial E}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial E}{\partial \mu} \quad \text{and} \quad 3D + r\frac{\partial D}{\partial r} + E$$

are even and odd functions respectively of μ and r. This in turn requires that D and E are (identical) odd functions of μr, for using (2·11) or (2·12) the even part of D, say D_e, must obey the equation

$$2D_e + r\frac{\partial D_e}{\partial r} = 0, \quad \text{i.e.} \quad D_e r^2 = \text{const.},$$

which is only consistent with regularity of D if $D_e = 0$.

Consequently, a solenoidal second-order axisymmetric tensor involving only similar vector quantities is symmetrical in its indices, and therefore has vanishing divergence with respect to both indices, and contains four arbitrary functions which are related by the two equations

$$4A + r\frac{\partial A}{\partial r} + \frac{1}{r}\frac{\partial B}{\partial r} - \frac{\mu}{r^2}\frac{\partial B}{\partial \mu} + \mu\frac{\partial D}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial D}{\partial \mu} = 0, \quad (2{\cdot}14)$$

$$\frac{1}{r}\frac{\partial B}{\partial \mu} + \mu\frac{\partial C}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial C}{\partial \mu} + 4D + r\frac{\partial D}{\partial r} = 0. \quad (2{\cdot}15)$$

For solenoidal second-order tensors which do involve scalars it is not possible to prove symmetry of the indices, and this has an important effect on the succeeding analysis for the following reason. The tensor obtained by differentiation of the solenoidal first-order tensor L_i is a second-order tensor $\partial L_i/\partial \xi_j$ which is still solenoidal in the index i. If this second-order tensor were symmetrical in its indices, then it would also be solenoidal in the index j, and hence

$$\nabla^2 L_i = 0.$$

Green's theorem then shows that if L_i vanishes for very large values of r, it is identically zero. Thus it may be stated that if all second-order axisymmetric tensors are symmetrical in their indices, then all first-order solenoidal axisymmetric tensors which vanish for large r are identically zero. However, a first-order tensor which vanishes for large r will usually be the correlation between a scalar at P, say, and a vector at P', for example, $\overline{pu'_i}$, so that in this case $\partial L_i/\partial \xi_j$ does involve a scalar term, rendering invalid the previous proof of symmetry in the indices of second-order tensors.

Trilinear form

The general third-order axisymmetric tensor T_{ijk} can readily be found to consist of the sum of terms $P\xi_i\xi_j\xi_k$, $Q\lambda_i\lambda_j\lambda_k$ and four groups of three, of which typical members are $R\xi_i\xi_j\lambda_k$, $S\xi_i\lambda_j\lambda_k$, $T\xi_i\delta_{jk}$ and $U\lambda_i\delta_{jk}$, where the arbitrary coefficients P, Q, etc., are functions of r^2 and μr. Third-order tensors play only a small direct part in the theory, so it will not be necessary to develop the details. The number of arbitrary functions is 14, and if the tensor is symmetrical in two of the indices, this number is reduced to 10. The divergence of T_{ijk} will be a second-order tensor of the form (2·8), so that if T_{ijk} is solenoidal in any of the indices there are five differential equations connecting the arbitrary functions.

3. THE FUNDAMENTAL VELOCITY CORRELATION

The correlation* between instantaneous velocity components u_i and u'_j at two different points P and P' is basic to any practical theory of turbulence. Write

$$\overline{u_i u'_j} = R_{ij},$$

and since R_{ij} is a second-order axisymmetric tensor, it will be of the form

$$R_{ij} = A\xi_i\xi_j + B\delta_{ij} + C\lambda_i\lambda_j + D\lambda_i\xi_j + E\xi_i\lambda_j. \tag{3·1}$$

Comparing with known results for isotropic turbulence, it is seen that in this case C, D, E are zero and A and B are replaced by $\left(\frac{f-g}{r^2}\right)\overline{u^2}$ and $g\overline{u^2}$ respectively, where f and g are the functions introduced by de Karman & Howarth and $\overline{u^2}$ is the mean square velocity fluctuation.

In virtue of continuity,
$$\frac{\partial R_{ij}}{\partial \xi_j} = \frac{\partial \overline{u_i u'_j}}{\partial x'_j} = 0, \tag{3·2}$$

so that R_{ij} is a solenoidal tensor. Hence, since it consists only of similar vector quantities, it is symmetrical in the indices, giving $D = E$, and the relations (2·14) and (2·15) are valid.

Now when r is zero, P and P' coinciding, R_{ij} must degenerate to the form of a one-point tensor depending on the vectors $\boldsymbol{\lambda}$, $\mathbf{a}$ and $\mathbf{b}$, i.e. R_{ij} becomes

$$R_{ij} = B_0\delta_{ij} + C_0\lambda_i\lambda_j, \tag{3·3}$$

---

* Throughout this paper the term 'correlation' refers to the mean product of the quantities concerned, while the 'correlation coefficient' is obtained by dividing by the root-mean-squares of all the quantities occurring in the correlation.

where B_0 and C_0 are the values of B and C at $r=0$. Putting $i=j$ (but without summing), then

$$\overline{u_i^2} = B_0 + C_0 \lambda_i^2. \tag{3·4}$$

If $\overline{u_1^2}$ is the mean square of the velocity fluctuation parallel to $\boldsymbol{\lambda}$, and $\overline{u_2^2}$ is the mean square fluctuation in *any* direction perpendicular to $\boldsymbol{\lambda}$, then

$$B_0 = \overline{u_2^2}, \quad C_0 = \overline{u_1^2} - \overline{u_2^2}. \tag{3·5}$$

It will be useful to determine A, B, C and D for small values of r. Each of the functions being obviously regular, one may expand in powers of r, remembering that A, B and C are functions of $\mu^2 r^2$ and r^2 while D is a function of μr and r^2, being odd in μr. Thus

$$\left.\begin{aligned} A &= A_0 + \frac{1}{2!} r^2 A_2 + \dots, \\ B &= B_0 + \frac{1}{2!} r^2 B_2 + \dots, \\ C &= C_0 + \frac{1}{2!} r^2 C_2 + \dots, \\ D &= \mu r\left(D_0 + \frac{1}{2!} r^2 D_2 + \dots\right). \end{aligned}\right\} \tag{3·6}$$

The first term of each series is constant,

$$A_0 = a, \quad B_0 = \overline{u_2^2}, \quad C_0 = \overline{u_1^2} - \overline{u_2^2}, \quad D_0 = d, \tag{3·7}$$

while all other coefficients are functions of μ^2 of degree no higher than the power of r^2 by which they are multiplied. Substituting the series (3·6) in the continuity relations (2·14) and (2·15) and retaining only the lowest power of r in each case,

$$4A_0 + B_2 - \tfrac{1}{2}\mu \frac{dB_2}{d\mu} + D_0 = 0, \tag{3·8}$$

$$\frac{1}{2}\frac{dB_2}{d\mu} + \mu C_2 + \tfrac{1}{2}(1-\mu^2)\frac{dC_2}{d\mu} + 5\mu D_0 = 0. \tag{3·9}$$

B_2 will contain two terms, one a constant, the other proportional to μ^2; (3·8) then requires

$$B_2 = -4a - d + b\mu^2, \tag{3·10}$$

where b is an arbitrary constant. C_2 will be of the same form, and from (3·9)

$$C_2 = -b - c - 5d + c\mu^2, \tag{3·11}$$

where c is another arbitrary constant.

One is now in a position to write down the fundamental velocity correlation tensor as far as the second power of r, viz.

$$\begin{aligned} R_{ij} = a\xi_i\xi_j &+ [\overline{u_2^2} + \tfrac{1}{2}r^2(-4a - d + b\mu^2)]\,\delta_{ij} \\ &+ [(\overline{u_1^2} - \overline{u_2^2}) + \tfrac{1}{2}r^2(-b - c - 5d + c\mu^2)]\,\lambda_i\lambda_j + d\mu r(\lambda_i\xi_j + \xi_i\lambda_j). \end{aligned} \tag{3·12}$$

Just as in the theory of isotropic turbulence use is made of the measured curvature of the velocity correlation curve at $r = 0$, so with axisymmetric turbulence it is possible to measure combinations of the quantities a, b, c and d. The correlations most suitable for practical measurement are those for which

$$i = j, \quad \lambda_i = 1,$$

i.e. correlations between velocities parallel to the axis of symmetry, or the direction of mean flow. In this case (3·12) becomes

$$\overline{u_i'u_i} = a\xi_i^2 + \overline{u_1^2} + \tfrac{1}{2}r^2(-4a-b-c-6d+b\mu^2+c\mu^2)+2d\mu r\xi_i \quad \text{(no summation).} \tag{3·13}$$

The correlation between points spaced longitudinally is obtained by putting $\xi_i = r$, $\mu = 1$, giving

$$\overline{u_i'u_i} = \overline{u_1^2} + \tfrac{1}{2}r^2(-2a-2d), \tag{3·14}$$

while that between points spaced laterally is found by putting $\xi_i = 0$, $\mu = 0$, i.e.

$$\overline{u_i'u_i} = \overline{u_1^2} + \tfrac{1}{2}r^2(-4a-b-c-6d). \tag{3·15}$$

By analogy with the notation of the theory of isotropic turbulence, one might define

$$a = \frac{(\overline{u_1^2}.\overline{u_2^2})^{\frac{1}{2}}}{2\lambda_1^2}, \quad b = \frac{(\overline{u_1^2}.\overline{u_2^2})^{\frac{1}{2}}}{\lambda_2^2}, \quad c = \frac{(\overline{u_1^2}.\overline{u_2^2})^{\frac{1}{2}}}{\lambda_3^2}, \quad d = \frac{(\overline{u_1^2}.\overline{u_2^2})^{\frac{1}{2}}}{\lambda_4^2}, \tag{3·16}$$

such that $\lambda_1 \to \lambda$ (Taylor 1935), and $\lambda_2, \lambda_3, \lambda_4$ become infinitely large as the turbulence approaches the isotropic condition. λ_1, λ_2, λ_3 and λ_4 all have the dimensions of length, and may be unreal since one cannot predict *a priori* the sign of the quantities a, b, c and d.

4. SCALES OF AXISYMMETRIC TURBULENCE

It is equally evident from the foregoing section that various scales of turbulence can be defined, combinations of which can be determined experimentally. As before, the velocity correlation tensor is

$$R_{ij} = A\xi_i\xi_j + B\delta_{ij} + C\lambda_i\lambda_j + D(\lambda_i\xi_j + \xi_i\lambda_j), \tag{4·1}$$

and hence

$$\left.\begin{aligned} L_A &= (\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty r^2 A\,dr, & L_C &= (\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty C\,dr, \\ L_B &= (\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty B\,dr, & L_D &= (\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty rD\,dr, \end{aligned}\right\} \tag{4·2}$$

all of which have the dimensions of (length). The factor $(\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}$ is chosen arbitrarily, and appears to be the least unsuitable of the available factors having the appropriate dimensions. For these integrals to exist, r^2A, B, C and rD are assumed to become zero at some finite value of r. L_A, L_B, L_C and μL_D are all functions of μ^2. Two relations between these 'scales' are derived by integrating the continuity relations (2·14) and (2·15) after multiplication of the former by r^2 and the latter by r, viz.

$$L_A - L_B - \mu\frac{dL_B}{d\mu} - 2\mu L_D + (1-\mu^2)\frac{dL_D}{d\mu} = 0, \qquad (4·3)$$

$$\frac{dL_B}{d\mu} - \mu L_C + (1-\mu^2)\frac{dL_C}{d\mu} + 2L_D = 0. \qquad (4·4)$$

The measurable scales of turbulence are now defined by

$$(\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty R_{ij}\,dr.$$

Taking, for illustration, only the cases of scales formed from correlations between velocities in the direction of the axis of symmetry, i.e. for which $i = j$, $\lambda_i = 1$,

$$\begin{aligned}(\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty R_{ij}\,dr &= (\overline{u_1^2}.\overline{u_2^2})^{-\frac{1}{2}}\int_0^\infty (A\xi_i^2 + B + C + 2D\xi_i)\,dr \\ &= (L_A + L_B + L_C + 2L_D)_{\mu=1} \quad \text{when} \quad \xi_i = r, \qquad (4·5) \\ &= (L_B + L_C)_{\mu=0} \quad \text{when} \quad \xi_i = 0, \qquad (4·6)\end{aligned}$$

and some simplification may be attained by the use of (4·3) and (4·4) for particular values of μ.

5. DISSIPATION OF ENERGY

The classical expression for the rate of dissipation of the energy of turbulence in unit volume of the fluid due to viscosity is

$$W = \rho\nu\left\{2\overline{\left(\frac{\partial u}{\partial x}\right)^2} + 2\overline{\left(\frac{\partial v}{\partial y}\right)^2} + 2\overline{\left(\frac{\partial w}{\partial z}\right)^2} + \overline{\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)^2} + \overline{\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right)^2} + \overline{\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)^2}\right\}, \qquad (5·1)$$

where ν is the kinematic viscosity and u, v, w are components of turbulent velocity in the directions of the axes x, y, z. Now

$$\overline{\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}} = \left(\overline{\frac{\partial u'}{\partial x'}\frac{\partial v}{\partial y}}\right)_{r=0} = \left(\frac{\partial^2 \overline{u'v}}{\partial x'\partial y}\right)_{r=0} = \left(\frac{\partial^2 \overline{u'v}}{\partial x\partial y'}\right)_{r=0} = \overline{\frac{\partial u}{\partial y}\frac{\partial v}{\partial x}}, \qquad (5·2)$$

since $\overline{u'v}$ depends only on differences of co-ordinates, $x'-x$, $y'-y$, provided the turbulence is homogeneous. Equation (5·1) can thus be written

$$\begin{aligned}W &= \rho\nu\left\{\overline{\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)^2} + \overline{\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right)^2} + \overline{\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right)^2}\right\} \\ &= \rho\nu\overline{\omega_i\omega_i}, \qquad (5·3)\end{aligned}$$

where ω_i is the component of vorticity in the direction of the i-axis, and the repeated index i implies summation.

This dissipation of energy due to viscosity is made manifest as a change in the kinetic energy of turbulence. Thus

$$\frac{d\overline{u_i u_i}}{dt} = -\frac{2}{\rho}W = -2\nu\overline{\omega_i\omega_i}. \qquad (5\cdot4)$$

It is also possible to express $\overline{\omega_i\omega_i}$ in terms of the constants a, b, c, d in the case of axisymmetric turbulence. The typical mean square component of vorticity at P is

$$\overline{\left(\frac{\partial u_j}{\partial x_k}-\frac{\partial u_k}{\partial x_j}\right)^2} = \left(2\frac{\partial^2 R_{jk}}{\partial\xi_l\,\partial\xi_m} - \frac{\partial^2 R_{jl}}{\partial\xi_k\,\partial\xi_m} - \frac{\partial^2 R_{km}}{\partial\xi_j\,\partial\xi_l}\right)_{r=0,\,l=j,\,m=k}, \qquad (5\cdot5)$$

where $j \neq k$. The mean square resultant vorticity is obtained by summing (5·5) over the three possible pairs of values of j and k. Hence,

$$\begin{aligned}\overline{\omega_i\omega_i} &= \left(\frac{\partial^2 R_{jk}}{\partial\xi_j\,\partial\xi_k} - \frac{\partial^2 R_{jj}}{\partial\xi_k\,\partial\xi_k}\right)_{r=0} \\ &= -\left(\frac{\partial^2 R_{jj}}{\partial\xi_k\,\partial\xi_k}\right)_{r=0} = -(\nabla^2 R_{jj})_{r=0}. \qquad (5\cdot6)\end{aligned}$$

The expression for $\nabla^2 R_{ij}$, which will be required again later, is found by differentiation of equation (4·1):

$$\begin{aligned}\nabla^2 R_{ij} = {}& \xi_i\xi_j\left(\nabla^2 A + \frac{4}{r}\frac{\partial A}{\partial r} - \frac{4\mu}{r^2}\frac{\partial A}{\partial\mu}\right) + \delta_{ij}(\nabla^2 B + 2A) \\ & + \lambda_i\lambda_j\left(\nabla^2 C + \frac{4}{r}\frac{\partial D}{\partial\mu}\right) + (\lambda_i\xi_j + \xi_i\lambda_j)\left(\nabla^2 D + \frac{2}{r}\frac{\partial D}{\partial r} - \frac{2\mu}{r^2}\frac{\partial D}{\partial\mu} + \frac{2}{r}\frac{\partial A}{\partial\mu}\right), \qquad (5\cdot7)\end{aligned}$$

and consequently

$$\nabla^2 R_{jj} = r^2\nabla^2 A + 4r\frac{\partial A}{\partial r} + 6A + 3\nabla^2 B + \nabla^2 C + 2\mu r\nabla^2 D + 4\mu\frac{\partial D}{\partial r}. \qquad (5\cdot8)$$

Now the expansions of A, B, C and D in powers of r sufficiently far to determine $\nabla^2 R_{jj}$ when r is zero are known, viz.

$$\left.\begin{aligned} A &= a, \\ B &= \overline{u_2^2} + \tfrac{1}{2}r^2(-4a - d + b\mu^2), \\ C &= \overline{u_1^2} - \overline{u_2^2} + \tfrac{1}{2}r^2(-b - c - 5d + c\mu^2), \\ D &= \mu r d, \end{aligned}\right\} \qquad (5\cdot9)$$

and substituting in (5·8), the value of $\overline{\omega_i\omega_i}$ is found to be

$$\overline{\omega_i\omega_i} = 30a + 2c + 20d. \qquad (5\cdot10)$$

Hence (5·4) can be rewritten as

$$\frac{d(\overline{u_1^2} + 2\overline{u_2^2})}{dt} = -2\nu(30a + 2c + 20d). \qquad (5\cdot11)^{*}$$

---

[*] Professor Goldstein tells me that he obtained this dissipation equation a few years ago in some unpublished work.

The similarity between (5·11) and Taylor's result for the decay of the velocity fluctuation in isotropic turbulence, for which c and d are zero, is evident.

The mean squares of the vorticity components are easily found from the expression (5·5) for the typical component. From the equation (4·1) for R_{ij}, the general second derivative of R_{ij} at $r = 0$ is

$$\left(\frac{\partial^2 R_{jk}}{\partial \xi_l \partial \xi_m}\right)_{r=0} = a(\delta_{km}\delta_{jl}+\delta_{jm}\delta_{kl}) + (-d-4a)\,\delta_{lm}\delta_{jk} + (-b-c-5d)\,\delta_{lm}\lambda_j\lambda_k$$
$$+ b\delta_{jk}\lambda_l\lambda_m + d(\delta_{km}\lambda_j+\delta_{jm}\lambda_k)\,\lambda_l + d(\delta_{kl}\lambda_j+\delta_{jl}\lambda_k)\,\lambda_m + c\lambda_j\lambda_k\lambda_l\lambda_m. \quad (5\cdot 12)$$

Substituting the appropriate forms of this expression in (5·5) then gives

$$10a + 2d + (c+7d)\,(\lambda_j\lambda_l + \lambda_k\lambda_m)_{l=j,\,m=k}. \quad (5\cdot 13)$$

The mean squares of the components of vorticity parallel and perpendicular to the vector $\boldsymbol{\lambda}$ are respectively

$$\overline{\omega_1^2} = 10a + 2d, \quad (5\cdot 14)$$

$$\overline{\omega_2^2} = 10a + c + 9d. \quad (5\cdot 15)$$

6. EQUATIONS OF MOTION

Considerations of the dissipation by viscosity are sufficient to determine the decay of kinetic energy, and consequently of the mean squares of the velocity components when the turbulence is isotropic, but for non-isotropic turbulence the rates of change of particular components of the velocity must be obtained from the equations of motion. The equations of motion at the point $P(x_i)$ can be written

$$\frac{\partial u_i}{\partial t} + u_k \frac{\partial u_i}{\partial x_k} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu\nabla^2 u_i, \quad (6\cdot 1)$$

where
$$\nabla^2 \equiv \frac{\partial^2}{\partial x_k \partial x_k}.$$

Multiplying by u'_j, the j-component of velocity at the point $P'(x'_i)$, and taking the time mean,

$$\overline{u'_j \frac{\partial u_i}{\partial t}} + \frac{\partial \overline{(u_i u_k u'_j)}}{\partial x_k} = -\frac{1}{\rho}\frac{\partial \overline{p u'_j}}{\partial x_i} + \nu\nabla^2 \overline{u_i u'_j}. \quad (6\cdot 2)$$

Using $\dfrac{\partial}{\partial x_i} = -\dfrac{\partial}{\partial \xi_i}$, this becomes

$$\overline{u'_j \frac{\partial u_i}{\partial t}} - \frac{\partial \overline{(u_i u_k u'_j)}}{\partial \xi_k} = \frac{1}{\rho}\frac{\partial \overline{p u'_j}}{\partial \xi_i} + \nu\nabla^2 \overline{u_i u'_j}. \quad (6\cdot 3)$$

A similar equation can be obtained from a consideration of the equations of motion at $P'(x'_i)$, viz.

$$\overline{u_i \frac{\partial u'_j}{\partial t}} + \frac{\partial \overline{(u_i u'_k u'_j)}}{\partial \xi_k} = -\frac{1}{\rho}\frac{\partial \overline{p' u_i}}{\partial \xi_j} + \nu\nabla^2 \overline{u_i u'_j}. \quad (6\cdot 4)$$

Adding (6·3) and (6·4),

$$\frac{\partial R_{ij}}{\partial t} - \frac{\partial}{\partial \xi_k}(\overline{u_i u_k u'_j} - \overline{u_i u'_k u'_j}) = \frac{1}{\rho}\left(\frac{\partial \overline{p u'_j}}{\partial \xi_i} - \frac{\partial \overline{p' u_i}}{\partial \xi_j}\right) + 2\nu \nabla^2 R_{ij}, \tag{6·5}$$

and this is the well-known fundamental equation for the change of R_{ij} with time.

Writing
$$T_{ij} = -\frac{\partial}{\partial \xi_k}(\overline{u_i u_k u'_j} - \overline{u_i u'_k u'_j}) \tag{6·6}$$

and
$$P_{ij} = \frac{1}{\rho}\left(\frac{\partial \overline{p u'_j}}{\partial \xi_i} - \frac{\partial \overline{p' u_i}}{\partial \xi_j}\right), \tag{6·7}$$

equation (6·5) appears in the concise form

$$\frac{\partial R_{ij}}{\partial t} + T_{ij} = P_{ij} + 2\nu \nabla^2 R_{ij}. \tag{6·8}$$

The velocity-pressure correlation

By reflecting the vector configuration associated with $\overline{p' u_i}$ in the point P, one obtains

$$\overline{p' u_i} = -(\overline{p u'_i})_{-\lambda}, \tag{6·9}$$

where the suffix indicates that the direction of $\boldsymbol{\lambda}$ is reversed. Hence

$$P_{ij} = \frac{1}{\rho}\frac{\partial \overline{p u'_j}}{\partial \xi_i} + \frac{1}{\rho}\frac{\partial (\overline{p u'_i})_{-\lambda}}{\partial \xi_j}. \tag{6·10}$$

P_{ij} can be given in terms of the scalar functions of the correlation between the pressure and velocity at two different points if one writes, in agreement with (2·4),

$$\frac{1}{\rho}\overline{p u'_i} = M\xi_i + N\lambda_i. \tag{6·11}$$

Substituting in (6·10),

$$\begin{aligned} P_{ij} = {} & \xi_i \xi_j \left(\frac{1}{r}\frac{\partial M}{\partial r} + \frac{1}{r}\frac{\partial M_-}{\partial r} - \frac{\mu}{r^2}\frac{\partial M}{\partial \mu} - \frac{\mu}{r^2}\frac{\partial M_-}{\partial \mu}\right) + \delta_{ij}(M + M_-) + \lambda_i \lambda_j \left(\frac{1}{r}\frac{\partial N}{\partial \mu} - \frac{1}{r}\frac{\partial N_-}{\partial \mu}\right) \\ & + \lambda_i \xi_j \left(\frac{1}{r}\frac{\partial M}{\partial \mu} - \frac{1}{r}\frac{\partial N_-}{\partial r} + \frac{\mu}{r^2}\frac{\partial N_-}{\partial \mu}\right) + \lambda_j \xi_i \left(\frac{1}{r}\frac{\partial M_-}{\partial \mu} + \frac{1}{r}\frac{\partial N}{\partial r} - \frac{\mu}{r^2}\frac{\partial N}{\partial \mu}\right), \end{aligned} \tag{6·12}$$

where
$$M_- = (M)_{-\lambda} \quad \text{and} \quad N_- = (N)_{-\lambda}.$$

Remembering that M and N are analytic functions of μr and r^2, one can expand in powers of r, thus

$$M = m_0 + \mu r m_1 + \tfrac{1}{2}r^2(m_2 + m'_2\mu^2) + \ldots, \tag{6·13}$$

$$N = n_0 + \mu r n_1 + \tfrac{1}{2}r^2(n_2 + n'_2\mu^2) + \tfrac{1}{6}\mu r^3(n_3 + n'_3\mu^2) + \ldots \tag{6·14}$$

Now the divergence of $\overline{p u'_i}$ is zero, so that M and N are connected by the relation

$$3M + r\frac{\partial M}{\partial r} + \mu\frac{\partial N}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial N}{\partial \mu} = 0,$$

and on substitution of the series (6·13) and (6·14), there is obtained

$$3m_0+n_1 = 0, \tag{6·15}$$

$$4m_1+n_2+n_2' = 0, \tag{6·16}$$

$$\tfrac{5}{2}m_2+\tfrac{1}{6}n_3 = 0, \tag{6·17}$$

$$\tfrac{5}{2}m_2'+\tfrac{1}{3}n_3+\tfrac{1}{2}n_3' = 0. \tag{6·18}$$

Substituting the expansions (6·13) and (6·14) in the expression (6·12) for P_{ij} and making use of the above continuity relations gives for the five scalar functions P_1, P_2, P_3, P_4, P_5 defining P_{ij} (analogous to the scalar functions A, B, C, D, E respectively, defining the general second order tensor as in (2·9)),

$$P_1 = 2m_2+\ldots, \tag{6·19}$$

$$P_2 = 2m_0+r^2(m_2+m_2'\mu^2)+\ldots, \tag{6·20}$$

$$P_3 = -6m_0+r^2(-5m_2+10m_2\mu^2-5m_2'\mu^2)+\ldots, \tag{6·21}$$

$$P_4 = m_1-n_2+\mu r(m_2'-5m_2)+\ldots, \tag{6·22}$$

$$P_5 = -m_1+n_2+\mu r(m_2'-5m_2)+\ldots. \tag{6·23}$$

The triple-velocity correlation

Let T_1, T_2, T_3, T_4, T_5 be the corresponding scalar functions which represent T_{ij}. T_{ij} vanishes with r since

$$\begin{aligned}(T_{ij})_{r=0} &= \overline{\frac{\partial u_i}{\partial x_k}u_k u_j}+\overline{\frac{\partial u_j}{\partial x_k}u_k u_i}\\ &= \frac{\partial\overline{(u_i u_k u_j)}}{\partial x_k}\\ &= 0,\end{aligned}$$

which result is valid for any kind of homogeneous turbulence. Consequently T_2 and T_3 vanish with r.

The triple-velocity and velocity-pressure correlation tensors are related, since (6·8) shows that

$$\frac{\partial T_{ij}}{\partial \xi_j} = \frac{\partial P_{ij}}{\partial \xi_j}, \quad \frac{\partial T_{ij}}{\partial \xi_i} = \frac{\partial P_{ij}}{\partial \xi_i}, \tag{6·24}$$

or, using (6·6) and (6·7),

$$\frac{\partial^2\overline{u_i u_k' u_j'}}{\partial \xi_j \partial \xi_k} = -\frac{1}{\rho}\nabla^2\overline{p' u_i}, \quad \frac{\partial^2\overline{u_i u_k u_j'}}{\partial \xi_i \partial \xi_k} = -\frac{1}{\rho}\nabla^2\overline{p u_j'}. \tag{6·25}$$

The relations (6·24) require that

$$T_4-P_4 = T_5-P_5. \tag{6·26}$$

Consideration can now be given to the scalar equations derivable from the basic tensor equation (6·8). All the terms of this equation are second-order axisymmetric tensors. The first and last terms are symmetrical, solenoidal in both indices, and are capable of representation in terms of four scalar functions of μ^2r^2 and r^2, between which two relations exist. The tensors T_{ij} and P_{ij}, on the other hand, are not symmetrical, or solenoidal in either index, and thus require five independent scalar functions of μr and r^2 for their representation. Equating to zero the coefficients of $\xi_i\xi_j$, δ_{ij}, etc., in (6·8) one obtains five scalar equations, the last two of which are identical in virtue of (6·26). The four independent equations are:

$$\begin{aligned}\frac{\partial A}{\partial t}+T_1 &= P_1+2\nu(\text{coeff. of } \xi_i\xi_j \text{ in } \nabla^2 R_{ij})\\ &= P_1+2\nu\left(\nabla^2 A+\frac{4}{r}\frac{\partial A}{\partial r}-\frac{4\mu}{r^2}\frac{\partial A}{\partial\mu}\right), \qquad (6\cdot27)\end{aligned}$$

$$\frac{\partial B}{\partial t}+T_2 = P_2+2\nu(\nabla^2 B+2A), \qquad (6\cdot28)$$

$$\frac{\partial C}{\partial t}+T_3 = P_3+2\nu\left(\nabla^2 C+\frac{4}{r}\frac{\partial D}{\partial\mu}\right), \qquad (6\cdot29)$$

$$\frac{\partial D}{\partial t}+T_4 = P_4+2\nu\left(\nabla^2 D+\frac{2}{r}\frac{\partial D}{\partial r}-\frac{2\mu}{r^2}\frac{\partial D}{\partial\mu}+\frac{2}{r}\frac{\partial A}{\partial\mu}\right). \qquad (6\cdot30)$$

7. RATE OF CHANGE OF THE VELOCITY COMPONENTS

It was shown by Karman & Howarth, and by Robertson, that the single scalar equation derived from the equation of motion (6·8), for isotropic turbulence, yields several auxiliary equations of physical significance. Expanding the scalar function in powers of r^2, and equating coefficients of similar powers of r^2, there is obtained, in turn, equations for the time rate of change of the mean square velocity fluctuation, of the mean square vorticity, of the mean square of some flow function of dimensions (velocity) × (length)^2, etc. A similar procedure can be applied, with partially similar results, to the four scalar equations (6·27)–(6·30).

The equations expressing the rates of change of the mean squares of the velocity components are obtained by putting $r = 0$ in the scalar equations (6·28) and (6·29). Remembering that T_2 and T_3 vanish with r, and making use of the expansions (5·9) for the functions A, B, C, D, and (6·19) for the functions P_1, P_2, P_3, P_4, P_5, expressions are obtained for $\frac{d\overline{u_2^2}}{dt}$ and $\frac{d}{dt}(\overline{u_1^2}-\overline{u_2^2})$, whence

$$\frac{d\overline{u_1^2}}{dt} = -4m_0+2\nu(-10a-2b-2c-14d), \qquad (7\cdot1)$$

$$\frac{d\overline{u_2^2}}{dt} = 2m_0+2\nu(-10a+b-3d), \qquad (7\cdot2)$$

and these equations are consistent with the expression (5·11) previously found for the viscous dissipation of kinetic energy.

Inasmuch as the rates of decay of the velocity components depend on a, b, c and d, i.e. on expressions related to the curvatures of various velocity correlation curves at the origin, these results are similar to those for isotropic turbulence. However, the rates of decay for axisymmetric turbulence also involve the term in m_0, related to the two-point velocity-pressure correlation, which vanishes under the condition of isotropy.

Viscous dissipation

Consider first the terms in (7·1) and (7·2) which represent changes due to viscosity. It is not difficult to show that both these terms are necessarily negative. Terms like $\overline{\left(\frac{\partial u_1}{\partial x_k}\right)^2}$ (no summation) are essentially positive, thus providing inequalities in a, b, c and d. The general positive term is

$$-\left(\frac{\partial^2 R_{ij}}{\partial \xi_k \partial \xi_l}\right)_{i=j,\ k=l,\ r=0.}$$

and using the expression (3·12) for R_{ij}, there is obtained

$$(d+4a)-2\delta_{ik}a+\lambda_i^2(b+c+5d)-\lambda_k^2 b-4\lambda_i\lambda_k\delta_{ik}d-\lambda_k^2\lambda_i^2 c \nless 0 \tag{7·3}$$

(repeated indices not summed). Giving i, k, λ_i and λ_k various values, a number of inequalities is obtained and by simple manipulation the quantities

$$(-10a-2b-2c-14d) \quad \text{and} \quad (-10a+b-3d)$$

are found to be negative. Neither of them can be zero without at least two of the mean squares of velocity derivatives also being zero; this is hardly possible, since it requires a zero velocity derivative at a given point for all time.

It is also pertinent to inquire which of the two viscous decay terms is greater, since the manner in which the relative values of $\overline{u_1^2}$ and $\overline{u_2^2}$ change with time is of great interest. Particularly is one concerned to detect any tendency toward isotropy of the turbulence, since there exists a general impression that in practice turbulence *does* become isotropic in sufficient time. It has already been seen that both $\overline{u_1^2}$ and $\overline{u_2^2}$ decrease due to viscosity; the question now is, to what extent (if at all) does the larger component decrease faster than the smaller? To establish that the velocity components tend to equality, it is not sufficient to show that

$$(-10a-2b-2c-14d)<(-10a+b-3d),$$

when $\overline{u_1^2}>\overline{u_2^2}$, and vice versa when $\overline{u_1^2}<\overline{u_2^2}$; this inequality would be obeyed if $\overline{u_1^2}$ and $\overline{u_2^2}$ merely decreased due to the action of viscosity at a rate proportional to $\overline{u_1^2}$ and $\overline{u_2^2}$ respectively. The absolute difference $|\overline{u_1^2}-\overline{u_2^2}|$ would in this case be decreasing, but only because all intensities are reduced proportionally, and equality of $\overline{u_1^2}$ and $\overline{u_2^2}$ would not be reached in a finite time. A more valid criterion for a tendency toward equality due to viscosity is obtained by comparing the values of $\frac{1}{\overline{u_1^2}}\frac{d\overline{u_1^2}}{dt}$ and $\frac{1}{\overline{u_2^2}}\frac{d\overline{u_2^2}}{dt}$, since the example quoted above of a turbulence field in which all intensities decreased proportionally would give equality of these two quantities. Comparison should therefore be made between

$$\frac{-10a-2b-2c-14d}{\overline{u_1^2}} \quad \text{and} \quad \frac{-10a+b-3d}{\overline{u_2^2}}.$$

Now it so happens that velocity correlations which contain these expressions explicitly can be constructed and are measurable. Thus when $i = j$, the value of R_{ij}, correct to the second power of r, is obtained from (3·12) as

$$R_{ii} = \overline{u_2^2} + (\overline{u_1^2} - \overline{u_2^2})\,\lambda_i^2 + \tfrac{1}{2}r^2\left[2a\left(\frac{\xi_i}{r}\right)^2 - 4a - d + b\mu^2 + (-b-c-5d+c\mu^2)\,\lambda_i^2 + 4d\mu\lambda_i\left(\frac{\xi_i}{r}\right)\right], \tag{7·4}$$

with no summation of the index. Consider first the case $\lambda_i = 1$. Noting that $\dfrac{\xi_i}{r} = \mu$ in this case, R_{ii} becomes

$$R_{ii} = \overline{u_1^2} + \tfrac{1}{2}r^2[-a(4-2\mu^2) - (b+c)\,(1-\mu^2) - d(6-4\mu^2)].$$

Putting $\mu^2 = \frac{1}{3}$, and writing R' for the correlation so obtained,

$$R' = \overline{u_1^2}\left[1 - \frac{1}{2}\left(\frac{r}{\sqrt{3}\,\lambda'}\right)^2\right],$$

where
$$\frac{1}{\lambda'^2} = \frac{10a+2b+2c+14d}{\overline{u_1^2}}. \tag{7·5}$$

Consider now the case $\lambda_i = 0$ for which (7·4) becomes

$$R_{ii} = \overline{u_2^2} + \tfrac{1}{2}r^2\left[2a\left(\frac{\xi_i}{r}\right)^2 - 4a - d + b\mu^2\right],$$

and choosing $\mu^2 = \left(\dfrac{\xi_i}{r}\right)^2 = \frac{1}{3}$ once again, the correlation

$$R'' = \overline{u_2^2}\left[1 - \frac{1}{2}\left(\frac{r}{\sqrt{3}\,\lambda''}\right)^2\right]$$

follows, where
$$\frac{1}{\lambda''^2} = \frac{10a-b+3d}{\overline{u_2^2}}. \tag{7·6}$$

In isotropic turbulence, there would have been obtained

$$R' = R'' = \overline{u^2}\left[1 - \frac{5}{2}\left(\frac{r}{\sqrt{3}\,\lambda}\right)^2\right],$$

which is consistent with the above expressions, since $a \to \overline{u^2}/2\lambda^2$ as isotropy is approached, λ being the usual velocity correlation parameter (Taylor 1935; de Karman & Howarth 1938). R' and R'' are drawn in figure 2 for the case

$$\mu = \frac{\xi_i}{r} = +\frac{1}{\sqrt{3}}.$$

The criterion given previously for a tendency toward equality of the velocity components due to viscosity is now seen to depend on the relative values of the radii of curvature at the origin of two correlation coefficient curves, both of which are capable of experimental measurement. Thus, when $\overline{u_1^2} > \overline{u_2^2}$, there is a tendency towards equality if

$$\frac{-10a-2b-2c-14d}{\overline{u_1^2}} < \frac{-10a+b-3d}{\overline{u_2^2}},$$

i.e. if

$$\lambda' < \lambda'', \tag{7·7}$$

and vice versa if $\overline{u_1^2} < \overline{u_2^2}$. There is no known reason to suppose that this criterion could be proved to be always obeyed; for initial values of a, b, c and d can be chosen quite arbitrarily. The theory provides information about the development of a, b, c and d with time, but the initial values are theoretically arbitrary and only limited in practice by what can be achieved with turbulence producing devices and stream distortion. Thus it may well be that the criterion (7·7) is obeyed or disobeyed at different periods of the history of the fluid; what is of importance is to determine whether (7·7) is obeyed when some time has elapsed and the effect of the initial conditions has largely disappeared. This is entirely a matter for experiment at the present stage in the development of the theory.

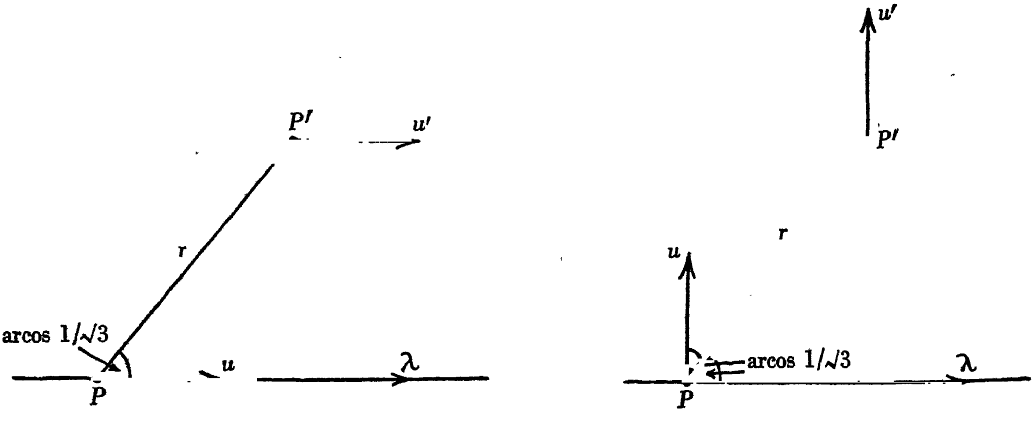


FIGURE 2

If the criterion (7·7) is found not to be obeyed, then it is natural to inquire if $|\overline{u_1^2} - \overline{u_2^2}|$ is decreasing. This inquiry can be answered by an examination of the correlation R_* given by (3·12) for the particular case $i \neq j$, $\xi_j = 0$, $\mu^2 = \frac{1}{3}$ and $\frac{\xi_i}{r} = \frac{2}{\sqrt{3}}\lambda_i$.

Thus

$$R_* = \lambda_i\lambda_j(\overline{u_1^2} - \overline{u_2^2})\left[1 + \tfrac{1}{2}r^2\,\frac{(-3b-2c-11d)}{3(\overline{u_1^2} - \overline{u_2^2})}\right]. \tag{7·8}$$

The sign of $(-3b-2c-11d)$, and hence of the viscous contribution to $\frac{d(\overline{u_1^2} - \overline{u_2^2})}{dt}$, is opposite to that of $(\overline{u_1^2} - \overline{u_2^2})$ if R_* falls off in absolute value with distance from the origin. A simpler correlation curve is therefore possible (requiring only one stationary point between the zero values at $\mp\infty$) if the difference between the velocity components tends to become smaller in absolute value. R_* could be measured directly, but a simpler method of measurement becomes evident when it is observed that

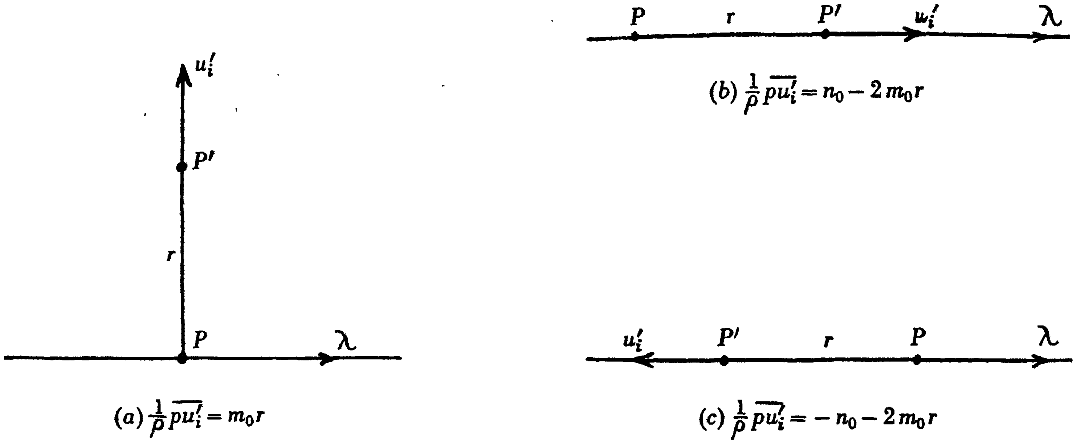
$$R_* = \lambda_i\lambda_j(R' - R''). \tag{7·9}$$

Transfer of energy

Consider now the velocity-pressure correlation term. This correlation contributes nothing to the decay of kinetic energy, which is effected solely by viscosity, but causes a transfer of energy between the lateral and longitudinal fluctuations. Now from (6·13), (6·14) and (6·15), the expression for $\frac{1}{\rho}\overline{pu_i'}$ when u_i' is parallel to $\boldsymbol{\xi}$ is

$$\frac{1}{\rho}\overline{pu_i'} = n_0\mu + m_0 r(1 - 3\mu^2) \tag{7·10}$$

as far as the first power of r, so that $\frac{1}{\rho}\overline{pu_i'} = m_0 r$ when $\mu = 0$, and $\mp n_0 - 2m_0 r$ when $\mu = \mp 1$.



(a) $\frac{1}{\rho}\overline{pu_i'} = m_0 r$

(b) $\frac{1}{\rho}\overline{pu_i'} = n_0 - 2m_0 r$

(c) $\frac{1}{\rho}\overline{pu_i'} = -n_0 - 2m_0 r$

FIGURE 3

These relations provide an interpretation of the mechanism of the transfer of energy. Suppose that m_0 is positive. Then the correlation shown in figure 3 (*a*) is positive, i.e. a positive pressure at the point P is associated, on the average, with a positive, or outward, lateral velocity at points displaced laterally from P. The pressure and the velocity at points close together are obviously related, the fluid always accelerating in directions of large negative pressure gradients, so that the outward velocity u_i' is in part a *result* of the positive pressure at P. The positive pressure at P can only have been produced as a result, in the mean, of a reduction in longitudinal velocities toward P. Thus when m_0 is positive, there is, in the mean, an addition to lateral fluctuations and a reduction in the magnitude of longitudinal fluctuations. The same deductions might have been made from the fact that the mean of the correlations shown in figures 3 (*b*) and (*c*) is negative when m_0 is positive, positive pressures at P being associated with longitudinal velocities toward P. The converse case of m_0 negative leads to the opposite conclusion, that there will be, in the mean, a reduction in the lateral fluctuation and an increase in the longitudinal fluctuations. Both these predictions are in accordance with the equations (7·1) and (7·2).

It is thus possible to understand why it is that when m_0 is positive, the longitudinal component supplies energy to the lateral, and vice versa, but the sign of m_0 has yet to be determined. Suppose that the longitudinal velocity fluctuation is greater than the lateral fluctuation. Then when positive pressures appear in the fluid at a point, they will have been produced, on the average, by the longitudinal fluctuation; the correlation between the pressure at P and the longitudinal velocity toward P at a point P' on the longitudinal axis through P is more likely to be positive than the correlation between the pressure at P and the inward lateral velocity at a neighbouring point P' laterally displaced. One of these correlations must be negative and the other positive. Hence when $\overline{u_1^2} > \overline{u_2^2}$, the mean of the correlations shown in figures 3 (*b*) and (*c*) is negative, i.e. $m_0 > 0$, and $\overline{u_2^2}$ increases with time at the expense of $\overline{u_1^2}$; the reverse is true when $\overline{u_2^2} > \overline{u_1^2}$. The effect of the velocity-pressure correlation is thus to transfer energy from the larger to the smaller of the velocity components.

To recapitulate the work of this section, the time rates of change of the longitudinal and lateral velocity fluctuations each contain two terms, one representing viscous decay, which is always negative and depends on the curvature of a certain velocity correlation curve near $r = 0$, and another depending on the correlation between the pressure and velocity at neighbouring points. Any tendency toward equality of $\overline{u_1^2}$ and $\overline{u_2^2}$ due to the viscous terms can be entirely determined from measurements of the velocity correlations R' and R''. If $R' - R''$ is found to fall off in absolute value with distance r from the origin (as is probable, since the contrary would require R_* to have at least three stationary points), then the absolute value of $\overline{u_1^2} - \overline{u_2^2}$ must be decreasing with time. Furthermore, if the radius of curvature at the origin ($r = 0$) of $R'/\overline{u_1^2}$ is less than that of $R''/\overline{u_2^2}$ when $\overline{u_1^2} > \overline{u_2^2}$, and vice versa when $\overline{u_1^2} < \overline{u_2^2}$, then $\overline{u_1^2}$ and $\overline{u_2^2}$ are being equalized quicker than would be the case if all mean-square velocities were being reduced at a rate proportional to their magnitude. The sign of the velocity-pressure correlation is such as to increase the energy of the smaller at the expense of that of the larger of the lateral and longitudinal components. The net result of these two effects is that there is always likely to be a tendency toward equality of $\overline{u_1^2}$ and $\overline{u_2^2}$, and although equality of $\overline{u_1^2}$ and $\overline{u_2^2}$ is not a sufficient condition for isotropy, it is probable that isotropy is attained with time. It should be possible to measure $d\overline{u_1^2}/dt$ and $d\overline{u_2^2}/dt$ in a practical experiment and also to determine the radii of curvature at the origin of the velocity correlations

R' and R'', giving the values of $(-10a-2b-2c-14d)$ and $(-10a+b-3d)$. Equations (7·1) and (7·2) then provide two independent estimates of m_0, the consistency of which would verify the analysis.

8. Rate of change of a, b, c, d

Since the four scalar equations (6·27)–(6·30) are true for all values of r, and all the terms are analytic functions of r, further equations may be obtained from a consideration of the coefficients of powers of r. Taking the coefficients of r^0 in (6·27), of r^2 in (6·28) and (6·29), and of r in (6·30), and making use of the expansions of R_{ij} and P_{ij} given in (3·6) and (6·19)–(6·23) respectively,

$$\frac{\partial a}{\partial t}+(T_1)_{r^0} = 2m_2+2\nu\left[7A_2-3\mu\frac{\partial A_2}{\partial \mu}+\tfrac{1}{2}(1-\mu^2)\frac{\partial^2 A_2}{\partial \mu^2}\right], \tag{8·1}$$

$$\frac{1}{2}\frac{\partial}{\partial t}(-4a-d+b\mu^2)+(T_2)_{r^2} = (m_2+m_2'\mu^2)+\frac{\nu}{12}\left[20B_4-2\mu\frac{\partial B_4}{\partial \mu}+(1-\mu^2)\frac{\partial^2 B_4}{\partial \mu^2}+24A_2\right], \tag{8·2}$$

$$\frac{1}{2}\frac{\partial}{\partial t}(-b-c-5d+c\mu^2)+(T_3)_{r^2} = (-5m_2+10m_2\mu^2-5m_2'\mu^2)+\frac{\nu}{12}\left[20C_4-2\mu\frac{\partial C_4}{\partial \mu}+(1-\mu^2)\frac{\partial^2 C_4}{\partial \mu^2}+48\frac{\partial \mu D_2}{\partial \mu}\right], \tag{8·3}$$

$$\mu\frac{\partial d}{\partial t}+(T_4)_r = \mu(m_2'-5m_2)+2\nu\left[7\mu D_2+(1-3\mu^2)\frac{\partial D_2}{\partial \mu}+\tfrac{1}{2}\mu(1-\mu^2)\frac{\partial^2 D_2}{\partial \mu^2}+\frac{\partial A_2}{\partial \mu}\right], \tag{8·4}$$

where $(T_1)_{r^0}$ is the coefficient of r^0 in T_1, etc. A_2 and D_2 will be functions of μ^2 of the first degree, while B_4 and C_4 are functions of μ^2 of the second degree, so that

$$A_2 = a_1+a_2\mu^2, \qquad D_2 = d_1+d_2\mu^2, \qquad B_4 = b_1+b_2\mu^2+b_3\mu^4, \qquad C_4 = c_1+c_2\mu^2+c_3\mu^4.$$

The ten constants $a_1, a_2, d_1, \ldots$ will be related by equations obtained by considering the second lowest powers of r in the continuity relations (2·14) and (2·15).

Equations (8·1) and (8·4) contain terms in μ^0 and equations (8·2) and (8·3) contain terms in both μ^0 and μ^2. This appears to give six independent equations, but it is not difficult to show that continuity relations and equations (6·24) eliminate two of them. The four independent equations giving the time rates of change of the factors a, b, c, d and, consequently, of the mean squares of the vorticity components, are

$$\frac{da}{dt}+(T_1)_{r^0} = 2m_2+2\nu(7a_1+a_2), \tag{8·5}$$

$$\frac{1}{2}\frac{db}{dt}+(T_2)_{r^2\mu^2} = m_2'+\frac{\nu}{12}(24a_2+14b_2+12b_3), \tag{8·6}$$

$$\frac{1}{2}\frac{dc}{dt}+(T_3)_{r^2\mu^2} = 10m_2-5m_2'+\frac{\nu}{12}(14c_2+12c_3+144d_2), \tag{8·7}$$

$$\frac{dd}{dt}+(T_4)_{r\mu} = m_2'-5m_2+2\nu(2a_2+7d_1+3d_2). \tag{8·8}$$

Using (5·14) and (5·15) the rates of change of the mean squares of the vorticity components are given by

$$\frac{d\overline{\omega_1^2}}{dt}+10(T_1)_{r^0}+2(T_4)_{r\mu} = 2m_2'+10m_2+2\nu(70a_1+14a_2+14d_1+6d_2), \tag{8·9}$$

$$\frac{d\overline{\omega_2^2}}{dt}+10(T_1)_{r^0}+2(T_3)_{r^2\mu^2}+9(T_4)_{r\mu} = -m_2'-5m_2+2\nu(70a_1+28a_2+\tfrac{7}{6}c_2+c_3+63d_1+39d_2). \tag{8·10}$$

The equations for the rate of change of vorticity can also be obtained directly from the equations of motion, although this direct procedure provides no information about da/dt, db/dt, dc/dt and dd/dt. After eliminating the pressure from the equations of motion, one obtains

$$\frac{\partial \omega_i}{\partial t}+u_k\frac{\partial \omega_i}{\partial x_k}-\omega_k\frac{\partial u_i}{\partial x_k} = \nu\nabla^2\omega_i. \tag{8·11}$$

Giving i the values 1 and 2 representing longitudinal and lateral quantities respectively, multiplying by ω_1 and ω_2, and taking mean values,

$$\frac{1}{2}\frac{d\overline{\omega_1^2}}{dt}+\overline{\tfrac{1}{2}u_k\frac{\partial \omega_1^2}{\partial x_k}}-\overline{\omega_1\omega_k\frac{\partial u_1}{\partial x_k}} = \nu\overline{\omega_1\nabla^2\omega_1}, \tag{8·12}$$

$$\frac{1}{2}\frac{d\overline{\omega_2^2}}{dt}+\overline{\tfrac{1}{2}u_k\frac{\partial \omega_2^2}{\partial x_k}}-\overline{\omega_2\omega_k\frac{\partial u_2}{\partial x_k}} = \nu\overline{\omega_2\nabla^2\omega_2}. \tag{8·13}$$

Equations (8·12) and (8·13) are identical with those given above, viz. (8·9) and (8·10), although this is not immediately clear due to the apparent presence in (8·9) and (8·10) of scalars of the velocity-pressure correlation. However, the first of equations (6·24) is

$$\frac{\partial(T_{ij}-P_{ij})}{\partial \xi_j} = 0,$$

and equating to zero the coefficient of ξ_i in this first order tensor, it follows that

$$\left(4+r\frac{\partial}{\partial r}\right)(T_1-P_1)+\left(\frac{1}{r}\frac{\partial}{\partial r}-\frac{\mu}{r^2}\frac{\partial}{\partial \mu}\right)(T_2-P_2)+\left(\mu\frac{\partial}{\partial r}+\frac{1-\mu^2}{r}\frac{\partial}{\partial \mu}\right)(T_5-P_5) = 0. \tag{8·14}$$

Putting both μ and r equal to zero gives

$$4(T_1)_{r^0}+2(T_2)_{r^2\mu^0}+(T_5)_{r\mu} = 4(P_1)_{r^0}+2(P_2)_{r^2\mu^0}+(P_5)_{r\mu} = m_2'+5m_2 \tag{8·15}$$

from (6·19)–(6·23), and consequently the factor $m_2'+5m_2$, which occurs in the vorticity decay equations (8·9) and (8·10), can be expressed in terms of scalars of the triple-velocity correlation. Thus the velocity-pressure correlation does not affect the decay of vorticity except inasmuch as the velocity-pressure and triple-velocity correlations are related by equations (6·24).

The identity of the pairs of equations (8·9) and (8·10), and (8·12) and (8·13) can now be established without difficulty. For instance,

$$10(T_1)_{r^0}+2(T_4)_{r\mu}-2(m_2'+5m_2) = 2(T_1)_{r^0}-4(T_2)_{r^2\mu^0} \quad \text{from (8·15) and (6·26),}$$

$$= \left(2\frac{\partial^2 T_{23}}{\partial\xi_2\partial\xi_3}-\frac{\partial^2 T_{22}}{\partial\xi_2^3}-\frac{\partial^2 T_{33}}{\partial\xi_2^2}\right)_{r=0},$$

where T_{23}, T_{22}, T_{33} are components of the tensor T_{ij},

$$= \overline{\omega_1\frac{\partial^2 u_3 u_k}{\partial x_2\partial x_k}}-\overline{\omega_1\frac{\partial^2 u_2 u_k}{\partial x_3\partial x_k}}$$

$$= \overline{\tfrac{1}{2}u_k\frac{\partial\omega_1^2}{\partial x_k}}-\overline{\omega_1\omega_k\frac{\partial u_1}{\partial x_k}},$$

in agreement with (8·12); with a similar proof

$$10(T_1)_{r^0}+2(T_3)_{r^2\mu^2}+9(T_4)_{r\mu}+(m_2'+5m_2) = \overline{\tfrac{1}{2}u_k\frac{\partial\omega_2^2}{\partial x_k}}-\overline{\omega_2\omega_k\frac{\partial u_2}{\partial x_k}}.$$

The identity of the remaining pairs of terms, i.e.

$$\overline{\omega_1\nabla^2\omega_1} = 2(70a_1+14a_2+14d_1+6d_2),$$

$$\overline{\omega_2\nabla^2\omega_2} = 2(70a_1+28a_2+\tfrac{7}{6}c_2+c_3+63d_1+39d_2),$$

is easily established by expressing $\overline{\omega_1\nabla^2\omega_1}$ and $\overline{\omega_2\nabla^2\omega_2}$ in terms of derivatives of the velocity correlation, and the details need not be exhibited here.

The interpretation of the terms in the equations (8·12) and (8·13) for the rate of decay of vorticity is similar to that for the corresponding equations relating to isotropic turbulence. Note first that the terms $\overline{u_k\frac{\partial\omega_1^2}{\partial x_k}}, = \frac{\partial}{\partial x_k}\overline{u_k\omega_1^2}$, and $\overline{u_k\frac{\partial\omega_2^2}{\partial x_k}}, = \frac{\partial}{\partial x_k}\overline{u_k\omega_2^2}$, are zero in virtue of the homogeneity of the turbulence. Also, using the suffix v to indicate a component in the direction of the vorticity vector,

$$\overline{\omega_1\omega_k\frac{\partial u_1}{\partial x_k}} = \overline{\omega^2\frac{\omega_1}{|\omega|}\frac{\partial u_1}{\partial x_v}}, \quad \overline{\omega_2\omega_k\frac{\partial u_2}{\partial x_k}} = \overline{\omega^2\frac{\omega_2}{|\omega|}\frac{\partial u_2}{\partial x_v}} \tag{8·16}$$

where $|\omega|$ is the magnitude of the vorticity vector. Hence equations (8·12) and (8·13) become

$$\frac{\partial\overline{\omega_1^2}}{\partial t} = 2\overline{\omega^2\frac{\omega_1}{|\omega|}\frac{\partial u_1}{\partial x_v}}+2\nu\overline{\omega_1\nabla^2\omega_1}, \tag{8·17}$$

$$\frac{\partial\overline{\omega_2^2}}{\partial t} = 2\overline{\omega^2\frac{\omega_2}{|\omega|}\frac{\partial u_2}{\partial x_v}}+2\nu\overline{\omega_2\nabla^2\omega_2}. \tag{8·18}$$

The first terms on the right-hand sides represent the correlation between the square of the total vorticity, and the spatial acceleration along the axis of the local vortex line due to the longitudinal or lateral velocity alone. These terms thus provide a measure of the extent to which vorticity is produced by the diffusive stretching of vortex filaments (Taylor 1938). The second terms on the right-hand sides represent the decrease of vorticity due to viscous dissipation.

Further equations for the rates of change of parameters defining the basic velocity correlation could be derived by considering higher powers of r in the equations (6·27)–(6·30), but it is clear from the deductions already made that their complexity would limit their practical value. All terms derived from the double-velocity correlation are measurable, but there have as yet been no measurements of either triple-velocity or velocity-pressure correlations, and these terms must therefore be deduced indirectly from the equations given.

REFERENCES

Batchelor, G. K. 1945 *Australian Council for Aeronautics, Report*, A.C.A. 13.
Karman, T. de & Howarth, L. 1938 *Proc. Roy. Soc.* A, **164**, 192.
Robertson, H. P. 1940 *Proc. Camb. Phil. Soc.* **36**, 209.
Taylor, G. I. 1935 *Proc. Roy. Soc.* A, **151**, 421.
Taylor, G. I. 1938 *Proc. Roy. Soc.* A, **164**, 15.

The correspondence between the particle and the wave aspects of the meson and the photon

By HARISH-CHANDRA

J. H. Bhabha Student, Tata Institute of Fundamental Research, Bombay

(*Communicated by H. J. Bhabha, F.R.S.—Received* 10 *September* 1945)

In the Kemmer formulation of the meson equation certain Γ_k^* matrices are introduced. They are determined completely by the condition that for a Lorentz transformation $\Gamma_k^*\psi$ transforms as a vector. The Γ^*'s, together with another matrix β (defined in terms of them), turn out to be very convenient for discussing the properties of the meson matrices β_k. With their help, a formula for the spur of a product of any number of β_k's is derived both for the 10- and the 5-row representations. Also they enable one to bring out, without making use of an explicit representation, the complete equivalence of the Kemmer formulation with the vector and scalar meson equations in the usual tensor form. Besides, the use of the Γ^*-matrices now permits the introduction of nuclear interaction in the Kemmer formulation in an elegant way. The extension of the 'particle' formulation to the case of a particle of zero rest-mass is discussed. Finally, a specially convenient representation of the β_k's is given both for the case of spin 1 and of spin 0. The Γ^*'s and the β are explicitly given for this particular representation.

1. Kemmer (1939) has given a formulation of the theory of the meson[‡] based on the following equation:

$$i\beta^k\partial_k\psi+\chi\psi = 0. \tag{1}$$

[‡] The term 'meson' is used in this paper to denote any particle with spin 1 or 0, having a non-vanishing rest-mass.

Here β^k $(k = 0, 1, 2, 3)$ are matrices acting upon ψ which itself is a one-column matrix. The β^k's satisfy the following commutation rules:

$$\beta^k\beta^l\beta^m + \beta^m\beta^l\beta^k = g^{kl}\beta^m + g^{ml}\beta^k, \tag{2}$$

g^{kl} being the usual metric tensor for flat space-time,

$$g_{kl} = 0, \quad k \neq l; \quad g_{00} = -g_{11} = -g_{22} = -g_{33} = 1.$$

It is well known that (1) is completely equivalent to the usual 'wave' formulation of the meson equations. However, until now this fact could be established only by making use of an explicit representation of the β^k's. A method of proving it directly will be given in this paper. It will be shown that the correspondence between the two formulations is indeed so close and complete that one can go over with ease from one formulation to the other at any stage of the calculation.

Further, it was not possible until now to introduce in (1) the interaction between the meson and the heavy particles in an elegant way. In his attempt to take into account this nuclear interaction Wilson (1940) had to make explicit use of the cumbersome representation of the β-matrices. This completely destroyed the beauty of the 'particle' formulation. However, as a result of the recognition of the close correspondence mentioned above the nuclear interaction can now be included in the matrix scheme without affecting its elegance.

In using the matrix method for the quantum-mechanical treatment of the physical processes involving mesons it is necessary to evaluate the spurs of the multiple products of the β_k's. The usual methods for their evaluation are extremely cumbersome (cf. Booth & Wilson 1940). In this paper a general formula for the spur of a product of any number of β_k's will be derived quite easily. It is hoped that it will be of use in future calculations.

It is well known that the Kemmer formulation can be applied only to a particle of non-vanishing rest-mass, i.e. $\chi \neq 0$. The derivation of the second order equation

$$\partial_k\partial^k\psi + \chi^2\psi = 0 \tag{3}$$

from (1) depends directly on the explicit assumption $\chi \neq 0$. It will be shown that by a slight modification the same formalism can be adapted to the case of zero rest-mass so that it can now be applied with equal elegance to the case of the photon or the electromagnetic field.

2. The Hermitian conjugate of (2) is

$$\beta^{\dagger k}\beta^{\dagger l}\beta^{\dagger m} + \beta^{\dagger m}\beta^{\dagger l}\beta^{\dagger k} = g^{kl}\beta^{\dagger m} + g^{ml}\beta^{\dagger k},$$

where $\beta^{\dagger k}$ is the Hermitian conjugate of β^k. Thus the $\beta^\dagger_k$'s satisfy the same commutation rules as the β_k's. Therefore every irreducible representation of the β_k's is equivalent to the corresponding one of the $\beta^\dagger_k$'s, because it has been shown by Kemmer that two inequivalent irreducible representations of the β_k's having the same number of rows and columns do not exist. Therefore there exists a matrix Λ such that

$$\Lambda\beta^k\Lambda^{-1} = \beta^{\dagger k}. \tag{4}$$

It follows from (4) (see Pauli 1936) that Λ can always be chosen to be Hermitian, i.e.

$$\Lambda^\dagger = \Lambda. \tag{5}$$

The Hermitian conjugate of (1) is

$$-i\partial_k \psi^\dagger \beta^{\dagger k} + \chi\psi^\dagger = 0,$$

so that from (4)

$$-i\partial_k \psi^\dagger \Lambda \beta^k + \chi\psi^\dagger\Lambda = 0.$$

Put

$$\psi^\dagger\Lambda = \psi^*, \tag{6}$$

then ψ^* satisfies the equation

$$-i\partial_k \psi^* \beta^k + \chi\psi^* = 0. \tag{7}$$

Let I_{mn} be the matrices representing the infinitesimal transformations of the Lorentz group, so that by an infinitesimal transformation

$$x'_k = x_k + \epsilon_{kl}x^l \quad (\epsilon_{kl} = -\epsilon_{lk}) \tag{8a}$$

ψ is transformed to

$$\psi' = (1+\tfrac{1}{2}\epsilon_{kl}I^{kl})\,\psi. \tag{8b}$$

From the relativistic invariance of (7) it follows that

$$\psi^{*\prime} = \psi^*(1-\tfrac{1}{2}\epsilon_{kl}I^{kl}), \tag{8c}$$

and from (8) and (6) one obtains

$$I^{\dagger kl}\Lambda + \Lambda I^{kl} = 0. \tag{9}$$

Besides, I^{kl} satisfy the following well-known commutation relations

$$[I^{kl}, I^{mn}] = -g^{km}I^{ln} + g^{lm}I^{kn} + g^{kn}I^{lm} - g^{ln}I^{km}, \tag{10a}$$

$$[\beta^k, I^{lm}] = g^{kl}\beta^m - g^{km}\beta^l, \tag{10b}$$

where $[A, B] = AB - BA$. It is well known that in the present case

$$I^{kl} = \beta^k\beta^l - \beta^l\beta^k, \tag{11}$$

and (9) and (10) are satisfied.

Introduce a one-column vector-matrix T_k such that $\psi^* T_k$ transforms as a vector, so that from (8c)

$$-I_{kl}T_m = g_{km}T_l - g_{lm}T_k. \tag{12}$$

Put

$$\Gamma^*_k = T^\dagger_k\Lambda; \tag{13}$$

therefore from (12) and (9)

$$\Gamma^*_k I_{lm} = g_{kl}\Gamma^*_m - g_{km}\Gamma^*_l. \tag{14}$$

From (12) and (14) it follows that

$$I_k{}^l T_l = 3T_k, \tag{15a}$$

$$\Gamma^{*k} I_k{}^l = 3\Gamma^{*l}; \tag{15b}$$

thus from (15a) and (14)

$$3\Gamma^*_k T_l = \Gamma^*_k I_l{}^m T_m = g_{kl}\Gamma^*_m T^m - \Gamma^*_l T_k,$$

which immediately gives
$$\Gamma^*_k\mathcal{T}_l-\Gamma^*_l\mathcal{T}_k = 0,$$
and therefore
$$\Gamma^*_k\mathcal{T}_l = \tfrac{1}{4}g_{kl}\Gamma^*_m\mathcal{T}^m. \tag{16}$$
Also from (12) and (14)
$$\tfrac{1}{2}I_{mn}I^{mn}\mathcal{T}_k = -3\mathcal{T}_k, \tag{17a}$$
$$\tfrac{1}{2}\Gamma^*_kI_{mn}I^{mn} = -3\Gamma^*_k, \tag{17b}$$
so that
$$\Gamma^*_k[\beta_l,\ \tfrac{1}{2}I_{rs}I^{rs}]\,\mathcal{T}_m = 0.$$
But from (10b)
$$\begin{aligned}[\beta_l, \tfrac{1}{2}I_{rs}I^{rs}] &= I_{ls}\beta^s+\beta^sI_{ls}\\ &= 3\beta_l+2\beta^sI_{ls};\end{aligned} \tag{18}$$
therefore
$$3\Gamma^*_k\beta_l\mathcal{T}_m-2g_{lm}\Gamma^*_k\beta^s\mathcal{T}_s+2\Gamma^*_k\beta_m\mathcal{T}_l = 0.$$
On contracting l, m one gets
$$\Gamma^*_k\beta_l\mathcal{T}^l = 0,$$
so that
$$3\Gamma^*_k\beta_l\mathcal{T}_m+2\Gamma^*_k\beta_m\mathcal{T}_l = 0.$$
On interchanging l, m and subtracting it is found that
$$\Gamma^*_k\beta_l\mathcal{T}_m = \Gamma^*_k\beta_m\mathcal{T}_l,$$
so that
$$\Gamma^*_k\beta_l\mathcal{T}_m = 0. \tag{19}$$

It will now be shown that for a given representation of the β_k's the choice of $\mathcal{T}_k$ is unique apart from a numerical factor. For this purpose transform (12) into the corresponding spinor equations

$$-I_{\mu\nu}\mathcal{T}_{\alpha\dot{\beta}} = \tfrac{1}{2}(\epsilon_{\mu\alpha}\mathcal{T}_{\nu\dot{\beta}}+\epsilon_{\nu\alpha}\mathcal{T}_{\mu\dot{\beta}}), \tag{20a}$$
$$-I_{\dot{\mu}\dot{\nu}}\mathcal{T}_{\alpha\dot{\beta}} = \tfrac{1}{2}(\epsilon_{\dot{\mu}\dot{\beta}}\mathcal{T}_{\alpha\dot{\nu}}+\epsilon_{\dot{\nu}\dot{\beta}}\mathcal{T}_{\alpha\dot{\mu}}). \tag{20b}$$

The Greek alphabets will be used to denote spinor indices, the Latin aphabets being reserved for tensor indices. Here

$$I_{\mu\nu} = \tfrac{1}{4}I_{kl}\sigma^k_{\mu\dot{\lambda}}\sigma^{l,\dot{\lambda}}{}_{\nu}, \tag{21a}$$
$$I_{\dot{\mu}\dot{\nu}} = \tfrac{1}{4}I_{kl}\sigma^k_{\dot{\mu}\lambda}\sigma^{l,\lambda}{}_{\dot{\nu}}, \tag{21b}$$
$$\mathcal{T}_{\alpha\dot{\beta}} = \mathcal{T}_k\sigma^k_{\alpha\dot{\beta}}. \tag{21c}$$

The σ-symbols have the usual meaning. $\epsilon^{\mu\nu}$ and $\epsilon_{\mu\nu}$ are the fundamental antisymmetric spinors for raising and lowering the indices, $\epsilon^{12} = \epsilon_{12} = 1$. Similarly, $\epsilon^{\dot{1}\dot{2}} = \epsilon_{\dot{1}\dot{2}} = 1$. Every finite irreducible representation of $I_{\mu\nu}$ is characterized by a positive number k which can assume only integral or half-integral values. The following relation holds for a particular irreducible representation

$$-\tfrac{1}{2}I^{\mu\nu}I_{\mu\nu} = k(k+1). \tag{22a}$$

Similarly, a finite irreducible representation of $I^{\dot{\mu}\dot{\nu}}$ is characterized by a positive number (integral or half-integral) given by

$$-\tfrac{1}{2}I^{\dot{\mu}\dot{\nu}}I_{\dot{\mu}\dot{\nu}} = l(l+1). \tag{22b}$$

As is well known (Dirac 1936; Fierz 1939) for any given values of k and l the following relations hold:

$$\left.\begin{aligned}
&I_{\mu\nu}(k) = u_{\mu}(k)\,v_{\nu}(k) - k\epsilon_{\mu\nu},\\
&I_{\dot\mu\dot\nu}(l) = u_{\dot\mu}(l)\,v_{\dot\nu}(l) - l\epsilon_{\dot\mu\dot\nu},\\
&u_{\mu}(k)\,u^{\mu}(k-\tfrac{1}{2}) = v_{\mu}(k-\tfrac{1}{2})\,v^{\mu}(k) = 0,\\
&u_{\mu}(k)\,v^{\mu}(k) = 2k, \quad v^{\mu}(k)\,u_{\mu}(k) = 2k+1,\\
&v^{\mu}(k+\tfrac{1}{2})\,u^{\nu}(k+\tfrac{1}{2}) - u^{\nu}(k)\,v^{\mu}(k) = \epsilon^{\nu\mu},\\
&u_{\dot\mu}(l)\,u^{\dot\mu}(l-\tfrac{1}{2}) = v_{\dot\mu}(l-\tfrac{1}{2})\,v^{\dot\mu}(l) = 0,\\
&u_{\dot\mu}(l)\,v^{\dot\mu}(l) = 2l, \quad v^{\dot\mu}(l)\,u_{\dot\mu}(l) = 2l+1,\\
&v^{\dot\mu}(l+\tfrac{1}{2})\,u^{\dot\nu}(l+\tfrac{1}{2}) - u^{\dot\nu}(l)\,v^{\dot\mu}(l) = \epsilon^{\dot\nu\dot\mu},
\end{aligned}\right\} \tag{23}$$

where $u_{\mu}(k)$ is a spinor matrix with $2k+1$ rows and $2k$ columns, while $v_{\mu}(k)$ has $2k$ rows and $2k+1$ columns. Similarly for $u_{\dot\mu}(l)$ and $v_{\dot\mu}(l)$. Corresponding to definite values of k and l split up $T_{\alpha\dot\beta}$ in $T_{\alpha\dot\beta}(k,l)$ (cf. Bhabha 1945). From (20) it is found that

$$-\tfrac{1}{2}I^{\mu\nu}I_{\mu\alpha}T_{\alpha\dot\beta} = \tfrac{3}{4}T_{\alpha\dot\beta}, \tag{24a}$$

$$-\tfrac{1}{2}I^{\dot\mu\dot\nu}I_{\dot\mu\dot\nu}T_{\alpha\dot\beta} = \tfrac{3}{4}T_{\alpha\dot\beta}, \tag{24b}$$

from which it follows that $T_{\alpha\dot\beta}(k,l) = 0$ unless

$$k(k+1) = \tfrac{3}{4}, \quad l(l+1) = \tfrac{3}{4},$$

i.e. unless $k = \frac{1}{2}$, $l = \frac{1}{2}$. Therefore (20) reduces to

$$-I_{\mu\nu}(\tfrac{1}{2})\,T_{\alpha\dot\beta} = \tfrac{1}{2}\epsilon_{\mu\alpha}T_{\nu\dot\beta} + \tfrac{1}{2}\epsilon_{\nu\alpha}T_{\mu\dot\beta}, \tag{25a}$$

$$-I_{\dot\mu\dot\nu}(\tfrac{1}{2})\,T_{\alpha\dot\beta} = \tfrac{1}{2}\epsilon_{\dot\mu\dot\beta}T_{\alpha\dot\nu} + \tfrac{1}{2}\epsilon_{\dot\nu\dot\beta}T_{\alpha\dot\mu}, \tag{25b}$$

where $T_{\alpha\dot\beta}$ is written instead of $T_{\alpha\dot\beta}(\frac{1}{2},\frac{1}{2})$ for brevity. Multiply (25a) by $v^{\mu}(\frac{1}{2})$ so that from (23)

$$-\tfrac{3}{2}v_{\nu}(\tfrac{1}{2})\,T_{\alpha\dot\beta} = \tfrac{1}{2}v_{\alpha}(\tfrac{1}{2})\,T_{\nu\dot\beta} + \tfrac{1}{2}\epsilon_{\nu\alpha}v^{\mu}(\tfrac{1}{2})\,T_{\mu\dot\beta}.$$

On interchanging ν, α and adding it is found that

$$v_{\nu}(\tfrac{1}{2})\,T_{\alpha\dot\beta} = -v_{\nu}(\tfrac{1}{2})\,T_{\nu\dot\beta},$$

so that

$$-v_{\nu}(\tfrac{1}{2})\,T_{\alpha\dot\beta} = \tfrac{1}{2}\epsilon_{\nu\alpha}v^{\mu}(\tfrac{1}{2})\,T_{\mu\dot\beta},$$

or, on multiplying by $u^{\nu}(\frac{1}{2})$,

$$T_{\alpha\dot\beta} = u_{\alpha}(\tfrac{1}{2})\,T_{\dot\beta}, \tag{26}$$

where $T_{\dot\beta} = \frac{1}{2}v^{\mu}(\frac{1}{2})\,T_{\mu\dot\beta}$. Similarly, by substituting (26) in (25b) one obtains

$$T_{\alpha\dot\beta}(\tfrac{1}{2},\tfrac{1}{2}) = u_{\alpha}(\tfrac{1}{2})\,u_{\dot\beta}(\tfrac{1}{2})\,T, \tag{27a}$$

where

$$T = \tfrac{1}{2}v^{\dot\beta}(\tfrac{1}{2})\,T_{\dot\beta} = \tfrac{1}{4}v^{\alpha}(\tfrac{1}{2})\,v^{\dot\beta}(\tfrac{1}{2})\,T_{\alpha\dot\beta}(\tfrac{1}{2},\tfrac{1}{2}). \tag{27b}$$

Similarly from (14)

$$\Gamma^{*}_{\alpha\dot\beta}(\tfrac{1}{2},\tfrac{1}{2}) = v_{\alpha}(\tfrac{1}{2})\,v_{\dot\beta}(\tfrac{1}{2})\,\Gamma^{*}, \tag{27c}$$

where

$$\Gamma^{*} = \tfrac{1}{4}\Gamma^{*}_{\alpha\dot\beta}(\tfrac{1}{2},\tfrac{1}{2})\,u^{\alpha}(\tfrac{1}{2})\,u^{\dot\beta}(\tfrac{1}{2}). \tag{27d}$$

In case‡ k, l, m, n form a complete system of commuting variables, T and Γ^* are ordinary numbers. That k, l, m, n do actually form a complete system of commuting variables for the various irreducible representations of the β_k's follows from the dimensionality of the 5- and 10-row representations (cf. § 6) which are the only non-trivial irreducible representations as shown by Kemmer. Therefore $T_{\alpha\dot{\beta}}$ and $\Gamma^*_{\alpha\dot{\beta}}$ are determined completely except for a numerical constant. Incidentally (27) proves that non-vanishing T_k and Γ^*_k satisfying (12) and (14) always exist whenever the irreducible representation $k = \frac{1}{2}$, $l = \frac{1}{2}$ is contained in the representation of the β_k's. This condition is of course satisfied for the case of spin 1 and 0 (cf. § 6).

Notice that from (27)

$$\Gamma^*_k T^k = \tfrac{1}{2}\Gamma^*_{\alpha\dot{\beta}} T^{\alpha\dot{\beta}} = \tfrac{1}{2}\Gamma^* T,$$

so that unless Γ^* or T is zero $\Gamma^*_k T^k \neq 0$. However, from (13) and the non-singularity of A it follows that if $T = 0$ then $\Gamma^* = 0$ and conversely. Therefore

$$\Gamma^*_k T^k \neq 0,$$

unless $\Gamma^*_k = T_k = 0$. Thus if T_k exists at all it can always be so normalized§ that

$$\Gamma^*_k T^k = 4.$$

This determines T_k completely except for a phase-factor of modulus unity. On introducing this normalization (16) becomes

$$\Gamma^*_k T_l = g_{kl}. \tag{28}$$

In view of (27), (19) is easy to understand. It has been shown by Bhabha (1945) that the only non-vanishing matrix elements of β_r are of the type

$$(k, l \,|\, \beta_r \,|\, k \pm \tfrac{1}{2}, l \pm \tfrac{1}{2}).$$

Since the only non-vanishing elements of Γ^*_k and T_l are $\Gamma^*_k(\frac{1}{2}, \frac{1}{2})$ and $T_l(\frac{1}{2}, \frac{1}{2})$, it follows that a product of any odd number of β_r's multiplied by Γ^*_k on the left and T_l on the right would give zero. Also notice that if θ be any operator commuting with I_{kl}, then θT_k and $\Gamma^*_k \theta$ satisfy equations entirely similar to (12) and (14). It therefore follows from the above result that θT_k and $\Gamma^*_k \theta$ can differ from T_k and Γ^*_k by a numerical factor only. One such operator θ is $\beta_m \beta^m$. Thus $\beta_m \beta^m T_k$ and $\Gamma^*_k \beta_m \beta^m$ are numerical multiples of T_k and Γ^*_k. Due to (13) the multiplying factors are conjugate complex of each other.

It may be mentioned here that all the results obtained so far depend only on the existence of I^{kl} and A satisfying (9) and (10) and are completely independent of the commutation rules (2). They hold therefore for the matrices β^l occurring in any relativistic wave equation of the type (1) (see Bhabha 1945). From now onwards (2) will be used explicitly.

---

‡ m and n are the eigenvalues of I^{12} and $I^{\dot{1}\dot{2}}$ respectively.

§ If necessary the sign of A may be reversed without disturbing (4) and (5) to make $\Gamma_k{}^* T^k$ positive.

Notice that from (2)
$$\beta_k\beta^k\beta_l+\beta_l\beta_k\beta^k = 5\beta_l, \tag{29a}$$

and
$$\beta^k\beta_l\beta_k = \beta_l, \tag{29b}$$

so that from (17) and (11)

$$\beta_m\beta_n(\beta^m\beta^n-\beta^n\beta^m)\mathcal{T}_k = \beta_n\beta^n(\beta_m\beta^m-4)\mathcal{T}_k = B(B-4)\mathcal{T}_k$$
$$= -3\mathcal{T}_k,$$

i.e.
$$(B-1)(B-3)\mathcal{T}_k = 0, \tag{30}$$

where B is a number given by

$$\beta^n\beta_n\mathcal{T}_k = B\mathcal{T}_k;$$

therefore $B = 1$ or $B = 3$. For the 10-row representation of the β_k's $\beta_m\beta^m$ has only 2 and 3 as eigenvalues, while for the 5-row representation the eigenvalues are 1 and 4.‡ Therefore for the 10-row representation $B = 3$ while for the 5-row one $B = 1$. Now consider $I_l^m I_{mn}\mathcal{T}_r$. From (12)

$$-I_l^m I_{mn}\mathcal{T}_r = I_{ln}\mathcal{T}_r - g_{nr}I_l^m\mathcal{T}_m$$
$$= -g_{ln}\mathcal{T}_r - 2g_{nr}\mathcal{T}_l.$$

On the other hand, from (11) and (29)

$$\begin{aligned} I_l^m I_{mn} &= (\beta^l\beta_m-\beta_m\beta_l)(\beta^m\beta_n-\beta_n\beta^m) \\ &= \beta_l\beta_m\beta^m\beta_n - 2\beta_l\beta_n + \beta_m\beta_l\beta_n\beta^m \\ &= \beta_l\beta_n(3-\beta_m\beta^m) + \beta_n\beta_l(1-\beta_m\beta^m) + g_{ln}\beta_m\beta^m, \end{aligned} \tag{31}$$

therefore
$$\{\beta_l\beta_n(3-B)+\beta_n\beta_l(1-B)+g_{ln}B\}\mathcal{T}_r = g_{ln}\mathcal{T}_r + 2g_{nr}\mathcal{T}_l. \tag{32}$$

For the 10-row representation $B = 3$ and so

$$\beta_n\beta_l\mathcal{T}_r = g_{nl}\mathcal{T}_r - g_{nr}\mathcal{T}_l. \tag{33a}$$

Multiplying (33a) by β^n

$$\beta^n\beta_n\beta_l\mathcal{T}_r = \beta_l(5-\beta^n\beta_n)\mathcal{T}_r = 2\beta_l\mathcal{T}_r = \beta_l\mathcal{T}_r - \beta_r\mathcal{T}_l,$$

i.e.
$$\beta_l\mathcal{T}_r = -\beta_r\mathcal{T}_l. \tag{33b}$$

---

‡ The easiest way to verify it is to calculate $\frac{1}{4}\beta_{\alpha\lambda}\beta^{\alpha\lambda}$ from the equation (100) of § 6. Notice that from (29a) it follows that $(\beta_m\beta^m-5/2)^2$ commutes with the β_k's and therefore for any irreducible representation of the β_k's it is a multiple of the unit matrix. Thus for any irreducible representation $\mathscr{B} = \beta_m\beta^m$ satisfies the equation

$$(\mathscr{B}-B_1)(\mathscr{B}-B_2) = 0,$$

where B_1, B_2 are numbers with $B_1+B_2 = 5$. Also it is easy to prove from (2) that the characteristic equation for $\mathscr{B}$ is

$$\mathscr{B}(\mathscr{B}-1)(\mathscr{B}-2)(\mathscr{B}-3)(\mathscr{B}-4) = 0;$$

this, together with (29a), permits only two possibilities

$$B_1 = 1, \quad B_2 = 4, \quad B_1 = 2, \quad B_2 = 3,$$

besides the trivial one $B = 0$ corresponding to $\beta_l = 0$.

Thus for spin 1, $\beta_l \mathcal{T}_r$ is antisymmetric in l, r.

On the other hand, for the 5-row representation $B = 1$, so that from (32)

$$\beta_l \beta_n \mathcal{T}_r = \mathcal{T}_l g_{nr}. \tag{34a}$$

Multiplying (34*a*) by β^l

$$\beta^l \beta_l \beta_n \mathcal{T}_r = \beta_n (5 - \beta^l \beta_l) \mathcal{T}_r = 4\beta_n \mathcal{T}_r = \beta^l \mathcal{T}_l g_{nr},$$

therefore

$$\beta_n \mathcal{T}_r = \tfrac{1}{4} g_{nr} \beta^l \mathcal{T}_l = \beta_r \mathcal{T}_n. \tag{34b}$$

The corresponding equations for Γ^*_k are

$$\left.\begin{aligned} \Gamma^*_k \beta_l \beta_m &= \Gamma^*_k g_{lm} - \Gamma^*_l g_{km} \\ \Gamma^*_k \beta_l &= -\Gamma^*_l \beta_k \end{aligned}\right\} \quad \text{for spin 1,} \tag{35}$$

$$\left.\begin{aligned} \Gamma^*_k \beta_l = \Gamma^*_l \beta_k &= \tfrac{1}{4} g_{kl} \Gamma^*_m \beta^m \\ \Gamma^*_k \beta_l \beta_m &= g_{kl} \Gamma^*_m \end{aligned}\right\} \quad \text{for spin 0.} \tag{36}$$

Now introduce a matrix β defined as follows:

$$\beta = \mathcal{T}_k \Gamma^{*k}. \tag{37}$$

From (28) it follows that

$$\beta^2 = \beta, \tag{38}$$

$$\Gamma^*_k \beta = \Gamma^*_k, \tag{39a}$$

$$\beta \mathcal{T}_k = \mathcal{T}_k; \tag{39b}$$

also from (33) and (35)

$$\beta_n \beta_l \beta = g_{nl} \beta - \mathcal{T}_l \Gamma^*_n = \beta \beta_n \beta_l \quad \text{for spin 1,}$$

and similarly from (34) and (36)

$$\beta_l \beta_n \beta = \mathcal{T}_l \Gamma^*_n = \beta \beta_l \beta_n \quad \text{for spin 0.}$$

Therefore in both cases

$$\beta_n \beta_l \beta - \beta \beta_n \beta_l = 0. \tag{40}$$

Now put

$$\beta \beta_k + \beta_k \beta = \theta_k; \tag{41}$$

then

$$\theta_k \theta_l = \beta_k \beta \beta_l + \beta \beta_k \beta_l \beta,$$

since $\beta \beta_k \beta = 0$ from (19). Also from (40) and (38)

$$\begin{aligned} \theta_k \theta_l \theta_m &= \beta_k \beta \beta_l \beta_m \beta + \beta \beta_k \beta_l \beta \beta_m \\ &= \beta_k \beta_l \beta_m \beta + \beta \beta_k \beta_l \beta_m, \end{aligned}$$

therefore

$$\begin{aligned} \theta_k \theta_l \theta_m + \theta_m \theta_l \theta_k &= (g_{kl} \beta_m + g_{ml} \beta_k) \beta + \beta (g_{kl} \beta_m + g_{ml} \beta_k) \\ &= g_{kl} \theta_m + g_{ml} \theta_k; \end{aligned} \tag{42}$$

thus the θ_k's satisfy the same commutation rules as the β_k's. Also from (19), (37) and (39)

$$\left.\begin{aligned} \theta_k \mathcal{T}_l &= (\beta \beta_k + \beta_k \beta) \mathcal{T}_l = \beta_k \mathcal{T}_l, \\ \Gamma^*_k \theta_l &= \Gamma^*_k \beta_l, \end{aligned}\right\} \tag{43a}$$

and
$$\left.\begin{aligned}\theta_k\theta_l\mathcal{T}_m &= \theta_k\beta_l\mathcal{T}_m = \beta\beta_k\beta_l\mathcal{T}_m = \beta_k\beta_l\beta\mathcal{T}_m = \beta_k\beta_l\mathcal{T}_m,\\ \Gamma^*_k\theta_l\theta_m &= \Gamma^*_k\beta_l\beta_m.\end{aligned}\right\} \tag{43b}$$

The representation of the θ_k's must be equivalent to that of the β_k's since from (43)

$$\beta_k\beta^k\mathcal{T}_l = B\mathcal{T}_l = \theta_k\theta^k\mathcal{T}_l,$$

and the B values for two inequivalent representations are different, therefore

$$\beta_k = S\theta_k S^{-1}, \tag{44}$$

where S is a non-singular matrix. Notice that from (41)

$$\theta_k\beta+\beta\theta_k = \theta_k, \tag{45}$$

so that from (44) and (45)
$$\beta_k\beta'+\beta'\beta_k = \beta_k, \tag{46}$$

where
$$\beta' = S\beta S^{-1} = S\mathcal{T}_k\Gamma^{*k}S^{-1}.$$

Put $\theta_k\theta_l-\theta_l\theta_k = J_{kl}$. Then from (43)

$$I_{kl}\mathcal{T}_m = J_{kl}\mathcal{T}_m = -g_{km}\mathcal{T}_l+g_{lm}\mathcal{T}_k.$$

But $J_{kl} = S^{-1}I_{kl}S$, and so

$$I_{kl}S\mathcal{T}_m = -g_{km}S\mathcal{T}_l+g_{lm}S\mathcal{T}_k.$$

Thus $S\mathcal{T}_k$ satisfies an equation similar to (12) and therefore must be equal to $c\mathcal{T}_k$, where c is an ordinary number. Similarly, $\Gamma^*_k S^{-1}$ is equal to $c^*\Gamma^*_k$, where c^* is another number. However,

$$cc^*\Gamma^*_k\mathcal{T}_l = \Gamma^*_k S^{-1}S\mathcal{T}_l = \Gamma^*_k\mathcal{T}_l = g_{kl},$$

therefore $cc^* = 1$ and

$$\beta' = S\mathcal{T}_k\Gamma^{*k}S^{-1} = cc^*\mathcal{T}_k\Gamma^{*k} = \mathcal{T}_k\Gamma^{*k} = \beta,$$

so that
$$\beta_k\beta+\beta\beta_k = \beta_k. \tag{47}$$

Collecting all the equations together

$$\beta = \mathcal{T}_k\Gamma^{*k}, \tag{48a}$$

$$\beta^2 = \beta, \tag{48b}$$

$$\beta\beta_k+\beta_k\beta = \beta_k, \tag{48c}$$

$$\Gamma^*_k\mathcal{T}_l = g_{kl}, \tag{48d}$$

$$\Gamma^*_k\beta_l\mathcal{T}_m = 0, \tag{48e}$$

$$\left.\begin{aligned}\beta_k\beta_l\mathcal{T}_m &= g_{kl}\mathcal{T}_m-g_{km}\mathcal{T}_l\\ \Gamma^*_k\beta_l\beta_m &= \Gamma^*_k g_{lm}-\Gamma^*_l g_{km}\\ \beta_k\mathcal{T}_l &= -\beta_l\mathcal{T}_k\\ \Gamma^*_k\beta_l &= -\Gamma^*_l\beta_k\end{aligned}\right\}\quad\text{for spin 1,} \tag{49a}$$

$$\left.\begin{aligned}\beta_k\beta_l\mathcal{T}_m &= \mathcal{T}_k g_{lm}\\ \Gamma^*_k\beta_l\beta_m &= g_{kl}\Gamma^*_m\\ \beta_l\mathcal{T}_m &= \tfrac{1}{4}g_{lm}\beta_k\mathcal{T}^k\\ \Gamma^*_k\beta_l &= \tfrac{1}{4}g_{kl}\Gamma^*_m\beta^m\end{aligned}\right\}\quad\text{for spin 0.} \tag{49b}$$

It will now be shown that the conditions (48*b*) and (48*c*) are so restrictive that for any irreducible set of matrices β_k (which need not in general satisfy (2)) there exist either none or just two matrices β satisfying them. Notice that if β is a solution then $1-\beta$ is also one. Further $\beta \neq 1-\beta$, since $\beta \neq \frac{1}{2}$ on account of (48*b*). Now let β' be any other solution. Then

$$\beta_k(\beta-\beta')+(\beta-\beta')\beta_k = 0,$$

so that $$[\beta_k, (\beta-\beta')^2] = [\beta_k, (\beta-\beta')]_+(\beta-\beta')-(\beta-\beta')[\beta_k, (\beta-\beta')]_+,$$

where $[A,B]_+ = AB+BA$. Since the β_k's are irreducible it follows that $(\beta-\beta')^2$ is a multiple of the unit matrix, i.e.

$$\beta^2+\beta'^2-\beta\beta'-\beta'\beta = \beta+\beta'-\beta\beta'-\beta'\beta = c, \tag{50}$$

where c is a number. On multiplying (50) by β or β'

$$\beta\beta'\beta = (1-c)\beta, \tag{51a}$$

$$\beta'\beta\beta' = (1-c)\beta'. \tag{51b}$$

From (48*c*) and (51) $$\beta\beta'\beta\beta_k+\beta_k\beta\beta'\beta = (1-c)\beta_k.$$

Now $$\begin{aligned}[\beta\beta'\beta, \beta_k]_+ &= \beta\beta'[\beta, \beta_k]_+ - \beta[\beta', \beta_k]_+\beta + [\beta, \beta_k]_+\beta'\beta \\ &= \beta\beta'\beta_k - \beta\beta_k\beta + \beta_k\beta'\beta \\ &= \beta\beta'\beta_k + \beta_k\beta'\beta = (1-c)\beta_k,\end{aligned} \tag{52a}$$

since $\beta\beta_k\beta = 0$ as follows from (48*c*). Similarly

$$[\beta'\beta\beta', \beta_k] = \beta'\beta\beta_k+\beta_k\beta\beta' = (1-c)\beta_k. \tag{52b}$$

Subtracting (52*b*) from (52*a*) one obtains

$$[(\beta'\beta-\beta\beta'), \beta_k] = 0,$$

therefore $\beta'\beta-\beta\beta'$ is a multiple of the unit matrix, i.e.

$$\beta'\beta-\beta\beta' = d, \tag{53}$$

where d is a number. Multiplying (53) by β from both sides

$$d\beta = 0,$$

therefore $d = 0$, since $\beta \neq 0$ from (48*c*). Thus

$$\beta'\beta = \beta\beta'. \tag{54}$$

Multiplying by β on the right gives

$$\beta'\beta = \beta\beta'\beta = (1-c)\beta,$$

while multiplication by β' on the left gives

$$\beta'\beta = \beta'\beta\beta' = (1-c)\beta',$$

therefore $$(1-c)\beta = (1-c)\beta'.$$

Thus either $\beta = \beta'$ or $c = 1$, in which case $\beta\beta' = \beta'\beta = 0$ and therefore

$$\beta' = 1-\beta$$

from (50).

Now consider the matrix $\gamma = \tfrac{1}{2}\beta_k \beta \beta^k$ (55)

for the case of spin 1:
$$\begin{aligned}\gamma^2 &= \tfrac{1}{4}\beta_k \beta \beta^k \beta^l \beta \beta_l = \tfrac{1}{4}\beta_k \beta^k \beta^l \beta \beta_l \\ &= \tfrac{1}{4}\beta^l (5-\beta_k \beta^k)\beta \beta_l \\ &= \tfrac{1}{4}\beta^l (5-B)\beta \beta_l \\ &= \tfrac{1}{2}\beta^l \beta \beta_l = \gamma.\end{aligned}$$

Further $\gamma = \tfrac{1}{2}\beta_k \beta \beta^k = \tfrac{1}{2}(1-\beta)\beta_k \beta^k = \tfrac{1}{2}(\beta_k \beta^k - 3\beta)$

therefore from (29*a*) and (48*c*)
$$\gamma \beta_k + \beta_k \gamma = \beta_k.$$

From the result just proved above it follows that either $\gamma = \beta$ or $\gamma = 1-\beta$. But
$$\begin{aligned}\beta^k \beta_k \gamma &= \beta^k \beta_k . \tfrac{1}{2}\beta_l \beta \beta^l \\ &= \tfrac{1}{2}\beta_l (5-\beta^k \beta_k)\beta \beta^l \\ &= 2\gamma,\end{aligned}$$

while
$$\beta^k \beta_k \beta = 3\beta,$$

therefore
$$\gamma = 1-\beta$$

and
$$\beta + \tfrac{1}{2}\beta^k \beta \beta_k = 1 \quad \text{for spin 1.} \tag{56}$$

For spin 0 take
$$\gamma = \tfrac{1}{4}\beta_k \beta \beta^k.$$

Then
$$\gamma^2 = \tfrac{1}{16}\beta^l (5-B)\beta \beta_l = \tfrac{1}{4}\beta^l \beta \beta_l = \gamma.$$

Also
$$\gamma = \tfrac{1}{4}\beta_k \beta \beta^k = \tfrac{1}{4}(1-\beta)\beta_k \beta^k = \tfrac{1}{4}(\beta_k \beta^k - \beta),$$

therefore
$$\gamma \beta_k + \beta_k \gamma = \beta_k$$

and $\gamma = \beta$ or $1-\beta$. But
$$\beta^l \beta_l \gamma = \tfrac{1}{4}\beta_k (5-\beta_l \beta^l)\beta \beta^l = 4\gamma,$$

while
$$\beta^l \beta_l \beta = \beta,$$

therefore
$$\gamma = 1-\beta$$

and
$$\beta + \tfrac{1}{4}\beta^k \beta \beta_k = 1 \quad \text{for spin 0.} \tag{57a}$$

Also from (49*b*) and (57*a*) it follows that
$$\beta^k \beta \beta^l = g^{kl}(1-\beta) \quad \text{for spin 0.} \tag{57b}$$

It is interesting to observe that for spin 0
$$\begin{aligned}\beta^k \beta^l \beta^m &= \beta^k \beta \beta^l \beta^m + \beta^k \beta^l \beta \beta^m \\ &= g^{kl}(1-\beta)\beta^m + \beta^k (1-\beta) g^{lm} \\ &= g^{kl}\beta^m \beta + \beta \beta^k g^{lm}.\end{aligned} \tag{58}$$

The condition (58) is stronger than (2) and may be used to separate the 5-row representation from the 10-row one.

A few other interesting results can also be derived. For example, consider‡
$$\omega = \frac{i}{4}\epsilon^{klmn}\beta_k \beta_l \beta_m \beta_n. \tag{59a}$$

---

‡ ϵ^{klmn} is a tensor antisymmetric in k, l, m, n and $\epsilon^{0123} = 1$.

It follows from (58) that $\omega = 0$ for spin 0. Therefore only the 10-row representation need be considered. In this case it is easy to prove from (48) and (49) that

$$\omega = \frac{i}{4}\epsilon^{klmn}\beta_k \mathcal{T}_l \Gamma^*_m \beta_n, \tag{59b}$$

so that

$$\begin{aligned} \omega^2 &= -\tfrac{1}{16}\epsilon^{klmn}\epsilon^{k'l'm'n'}\beta_k \mathcal{T}_l \Gamma^*_m \beta_n \beta_{k'} \mathcal{T}_{l'} \Gamma^*_{m'} \beta_{n'} \\ &= \tfrac{1}{8}\epsilon^{klmn}\epsilon_{mn}{}^{m'n'}\beta_k \mathcal{T}_l \Gamma^*_{m'} \beta_{n'} \\ &= -\tfrac{1}{2}\beta_k \mathcal{T}_l \Gamma^{*k}\beta^l = \tfrac{1}{2}\beta_k \mathcal{T}_l \Gamma^{*l}\beta^k \\ &= \tfrac{1}{2}\beta_k \beta \beta^k = 1-\beta. \end{aligned} \tag{59c}$$

It follows immediately from (59*b*) and (59*c*) that

$$\begin{aligned} \omega^3 &= \omega, \\ \omega^2\beta_k + \beta_k\omega &= \beta_k, \\ \omega\beta_r\omega &= 0, \\ \beta_r\omega\beta_s + \beta_s\omega\beta_r &= 0. \end{aligned}$$

Also from (59*b*)

$$\begin{aligned} \beta_r\beta_s\omega + \omega\beta_s\beta_r &= \frac{i}{2}\epsilon_s{}^{lmn}\beta_l(\mathcal{T}_m\Gamma^*_r - \mathcal{T}_r\Gamma^*_m)\,\beta_n \\ &= \frac{i}{2}\epsilon_s{}^{lmn}\beta_l\beta(\beta_m\beta_r - \beta_r\beta_m)\,\beta_n \\ &= \frac{i}{2}(1-\beta)\,\epsilon_s{}^{lmn}(\beta_l\beta_m\beta_r\beta_n - \beta_l\beta_r\beta_m\beta_n) \\ &= \frac{i}{4}(1-\beta)\,\epsilon_s{}^{lmn}(\beta_r\beta_l\beta_m\beta_n - \beta_l\beta_r\beta_m\beta_n + \beta_l\beta_m\beta_r\beta_n - \beta_l\beta_m\beta_n\beta_r) \quad \text{from (2)} \\ &= g_{sr}\omega^3 = g_{sr}\omega. \end{aligned}$$

Therefore, as has already been pointed out by Schrödinger (1943) and Kemmer (1943), ω satisfies the same commutation rules as any of the β_k's. Thus if one puts $\omega = \beta_4$ and $g_{44} = 1$, $g_{4k} = 0$ ($k \neq 4$), then (2) would hold for $k, l, m = (0, 1, 2, 3, 4)$.

3. The above results can now be applied to evaluate the spurs of any matrix composed of the β_k's and their products. Notice that

$$sp(\beta) = \Gamma^*_m \mathcal{T}^m = 4, \tag{60a}$$

$$sp(\beta_k) = sp(\beta\beta_k + \beta_k\beta) = 2\Gamma^*_m\beta_k\mathcal{T}^m = 0, \tag{60b}$$

and

$$sp(\beta_k\beta_l) = sp(\beta_k\beta\beta_l + \beta_k\beta_l\beta) = \Gamma^*_m\beta_l\beta_k\mathcal{T}^m + \Gamma^*_m\beta_k\beta_l\mathcal{T}^m,$$

so that from (49)

$$\left.\begin{aligned} sp(\beta_k\beta_l) &= 6g_{kl} \quad \text{for spin 1,} \\ sp(\beta_k\beta_l) &= 2g_{kl} \quad \text{for spin 0.} \end{aligned}\right\} \tag{60c}$$

Consider now the spin 1 case, so that

$$\Gamma^*_k\beta_l\beta_m\mathcal{T}_n = g_{kn}g_{lm} - g_{km}g_{ln} \equiv \eta_{[kl][mn]}. \tag{61a}$$

Due to (39)

$$\Gamma^*_{k_0}\beta_{k_1}\beta_{k_2}\ldots\beta_{k_n}\mathcal{T}_{k_{n+1}} = \Gamma^*_{k_0}\beta\beta_{k_1}\ldots\beta_{k_n}\beta\mathcal{T}_{k_{n+1}}.$$

In case n the number of factors $\beta_{k_1}, \ldots, \beta_{k_n}$ is odd one finds on using (40) that

$$\Gamma^*_{k_0}\beta_{k_1}\ldots\beta_{k_{2n-1}}\mathcal{T}_{k_{2n}} = \Gamma^*_{k_0}\beta_{k_1}\ldots\beta\beta_{k_{2n-1}}\beta\mathcal{T}_{k_{2n}} = 0. \tag{61b}$$

On the other hand, for an even number of factors one gets from (38), (39), (40), (37) and (61a)

$$\begin{aligned}\Gamma^*_{k_0}\beta_{k_1}\ldots\beta_{k_{2n}}\mathcal{T}_{k_{2n+1}} &= \Gamma^*_{k_0}\beta_{k_1}\beta_{k_2}\beta\beta_{k_3}\beta_{k_4}\beta\ldots\beta\beta_{k_{2n-1}}\beta_{k_{2n}}\mathcal{T}_{k_{2n+1}}\\ &= \eta_{[k_0 k_1][k_2}{}^{t_1]}\,\eta_{[t_1 k_3][k_4}{}^{t_2]}\ldots\eta_{[t_{n-1}k_{2n-1}][k_{2n}k_{2n+1}]}.\end{aligned} \tag{61c}$$

Now
$$\begin{aligned}sp(\beta_{k_1}\ldots\beta_{k_n}) &= sp(\beta\beta_{k_1}\ldots\beta_{k_n}+\beta_{k_1}\beta\beta_{k_2}\ldots\beta_{k_n})\\ &= \Gamma^*_m\beta_{k_1}\ldots\beta_{k_n}\mathcal{T}^m+\Gamma^*_m\beta_{k_2}\ldots\beta_{k_n}\beta_{k_1}\mathcal{T}^m,\end{aligned}$$

so that
$$\begin{aligned}sp(\beta_{k_1}\ldots\beta_{k_{2n}}) &= \eta_{[t_1 k_1][k_2}{}^{t_2]}\,\eta_{[t_2 k_3][k_4}{}^{t_3]}\ldots\eta_{[t_n k_{2n-1}][k_{2n}}{}^{t_1]}\\ &\quad+\eta_{[t_1 k_2][k_3}{}^{t_2]}\,\eta_{[t_2 k_4][k_5}{}^{t_3]}\ldots\eta_{[t_n k_{2n}][k_1}{}^{t_1]},\end{aligned} \tag{62a}$$

$$sp(\beta_1\ldots\beta_{k_{2n+1}}) = 0. \tag{62b}$$

For spin 0
$$\Gamma^*_k\beta_l\beta_m\mathcal{T}_n = g_{kl}g_{mn},$$

therefore
$$\Gamma^*_{k_0}\beta_{k_1}\ldots\beta_{k_{2n-1}}\mathcal{T}_{k_{2n}} = 0, \tag{63a}$$

$$\Gamma^*_{k_0}\beta_{k_1}\ldots\beta_{k_{2n}}\mathcal{T}_{k_{2n+1}} = g_{k_0k_1}g_{k_2k_3}\ldots g_{k_{2n}k_{2n+1}}. \tag{63b}$$

Thus
$$sp(\beta_{k_1}\ldots\beta_{k_{2n}}) = g_{k_1k_2}g_{k_3k_4}\ldots g_{k_{2n-1}k_{2n}}+g_{k_2k_3}g_{k_4k_5}\ldots g_{k_{2n}k_1}, \tag{64a}$$

$$sp(\beta_{k_1}\ldots\beta_{k_{2n+1}}) = 0. \tag{64b}$$

Multiplying (62b) and (64b) by arbitrary vectors $A^{(1)k_1}$, $A^{(2)k_2}$, …, $A^{(2n)k_{2n}}$, one finds that

$$\begin{aligned}&sp\{(\mathbf{A}^{(1)}\boldsymbol{\beta})(\mathbf{A}^{(2)}\boldsymbol{\beta})\ldots(\mathbf{A}^{(2n)}\boldsymbol{\beta})\}\\ &\quad= A^{(1,2)t_2}_{t_1}A^{(3,4)t_3}_{t_2}\ldots A^{(2n-1,2n)t_1}_{t_n}+A^{(2,3)t_2}_{t_1}A^{(4,5)t_3}_{t_2}\ldots A^{(2n,1)t_1}_{t_n}\quad\text{for spin 1,}\end{aligned} \tag{65a}$$

while for spin 0

$$\begin{aligned}&sp\{(\mathbf{A}^{(1)}\boldsymbol{\beta})(\mathbf{A}^{(2)}\boldsymbol{\beta})\ldots(\mathbf{A}^{(2n)}\boldsymbol{\beta})\}\\ &\quad= (\mathbf{A}^{(2)}\mathbf{A}^{(3)})(\mathbf{A}^{(4)}\mathbf{A}^{(5)})\ldots(\mathbf{A}^{(2n)}\mathbf{A}^{(1)})+(\mathbf{A}^{(1)}\mathbf{A}^{(2)})\ldots(\mathbf{A}^{(2n-1)}\mathbf{A}^{(2n)}),\end{aligned} \tag{65b}$$

where
$$(\mathbf{A}^{(r)}\boldsymbol{\beta}) = A^{(r)k}\beta_k,\quad (\mathbf{A}^{(r)}\mathbf{A}^{(s)}) = A^{(r)k}A^{(s)}_k,$$
$$A^{(r,s)l}_k = (\mathbf{A}^{(r)}\mathbf{A}^{(s)})\,\delta^l_k - A^{(r)l}A^{(s)}_k.$$

4. Let us now return to equation (1). First consider the spin 1 case. Multiplying (1) by Γ^*_l from the left

$$i\Gamma^*_l\beta_k\partial^k\psi+\chi\Gamma^*_l\psi = 0.$$

Similarly, multiplying by $\Gamma^*_k\beta_l$,

$$i\Gamma^*_k\beta_l\beta_m\partial^m\psi+\chi\Gamma^*_k\beta_l\psi = 0,$$

or from (49a)
$$i(\partial_l\Gamma^*_k\psi-\partial_k\Gamma^*_l\psi)+\chi\Gamma^*_k\beta_l\psi = 0.$$

Put
$$\Gamma^*_k\psi = U_k, \tag{66a}$$

$$\frac{\chi}{i}\Gamma^*_k\beta_l\psi = G_{kl}. \tag{66b}$$

On account of the antisymmetry of $\Gamma^*_k\beta_l$ in k, l, G_{kl} is also antisymmetric. The above equations can now be written as

$$\partial_k U_l - \partial_l U_k = G_{kl}, \tag{67a}$$

$$\partial^k G_{kl} + \chi^2 U_l = 0, \tag{67b}$$

which are precisely the usual equations of the wave formulation. To understand the equation

$$\partial_l \psi = \partial_k \beta^k \beta_l \psi \tag{68}$$

which follows from (1), replace β^k by $\beta\beta^k + \beta^k\beta$. Then

$$\begin{aligned}\beta^k\beta_l\psi &= \beta\beta^k\beta_l\psi + \beta^k\beta\beta_l\psi = \beta^k\beta_l\beta\psi + \beta^k\beta\beta_l\psi\\ &= \beta^k\beta_l\mathcal{T}_m\Gamma^{*m}\psi + \beta^k\mathcal{T}_m\Gamma^{*m}\beta_l\psi\\ &= (\delta^k_l\mathcal{T}_m - \delta^k_m\mathcal{T}_l)\,U^m + \frac{i}{\chi}\beta^k\mathcal{T}_m G^m{}_l.\end{aligned}$$

Similarly $$\psi = (\beta + \tfrac{1}{2}\beta^k\beta\beta_k)\,\psi = \mathcal{T}_m U^m + \tfrac{1}{2}\beta^k\mathcal{T}_m\frac{i}{\chi}G^m{}_k,$$
therefore from (68)

$$\mathcal{T}_k\partial_l U^k + \frac{i}{2\chi}\beta^k\mathcal{T}^m\partial_l G_{mk} = \mathcal{T}_k\partial_l U^k - \mathcal{T}_l\partial^k U_k + \frac{i}{\chi}\beta^k\mathcal{T}^m\partial_k G_{ml}. \tag{69}$$

Multiplying by Γ^{*r} one obtains $$\partial^k U_k = 0, \tag{70a}$$

while on multiplying by $\Gamma^{*r}\beta^s$ it is found that

$$\partial_l G_{rs} + \partial_s G_{lr} + \partial_r G_{sl} = 0. \tag{70b}$$

The equations (70) are completely equivalent to (69) or (68) because the latter follow from the former.

In the same way the energy-momentum tensor

$$T_{kl} = \chi\psi^*(g_{kl} - \beta_k\beta_l - \beta_l\beta_k)\,\psi \tag{71}$$

can also be transformed

$$\begin{aligned}\psi^*\beta_k\beta_l\psi &= \psi^*(\beta\beta_k\beta_l + \beta_k\beta\beta_l)\,\psi\\ &= \psi^*\mathcal{T}^m(\Gamma^*_m g_{kl} - \Gamma^*_k g_{ml})\,\psi + \psi^*\beta_k\mathcal{T}_m\Gamma^{*m}\beta_l\psi\\ &= g_{kl}U^{\dagger m}U_m - U^\dagger_l U_k - \frac{1}{\chi^2}G^\dagger_{km}G^m{}_l\end{aligned}$$

and $$\begin{aligned}\psi^*\psi &= \psi^*(\beta + \tfrac{1}{2}\beta^k\beta\beta_k)\,\psi\\ &= U^{\dagger m}U_m - \frac{1}{2\chi^2}G^{\dagger km}G_{mk} = U^{\dagger m}U_m + \frac{1}{2\chi^2}G^{\dagger mn}G_{mn},\end{aligned}$$

therefore

$$T_{kl} = \frac{1}{\chi}\{G^\dagger_{km}G^m{}_l + G^\dagger_{lm}G^m{}_k + \tfrac{1}{2}g_{kl}G^{\dagger mn}G_{mn} + \chi^2(U^\dagger_k U_l + U^\dagger_l U_k - g_{kl}U^{\dagger m}U_m).\} \tag{72}$$

The angular momentum tensor $M_{kl,m}$ is given by

$$M_{kl,m}=\frac{1}{2i}\left\{(\partial_l\psi^*x_k-\partial_k\psi^*x_l)\,\beta_m\psi-\psi^*(x_k\partial_l-x_l\partial_k)\,\beta_m\psi-\frac{1}{2i}\psi^*(I_{kl}\beta_m+\beta_m I_{kl})\,\psi\right\}. \tag{73}$$

It would be interesting to transform the spin term to the usual wave formulation. It is found that

$$\psi^*\beta_k\beta_l\beta_m\psi=\frac{i}{\chi}[G^{\dagger n}_k U_n g_{lm}-G^\dagger_{km}U_l-U^\dagger_l G_{km}+U^{\dagger n}G_{mn}g_{kl}],$$

therefore

$$\psi^*I_{kl}\beta_m\psi=\frac{i}{\chi}[G^{\dagger n}_k U_n g_{lm}-G^{\dagger n}_l U_n g_{km}-G^\dagger_{km}U_l+G^\dagger_{lm}U_k-U^\dagger_l G_{km}+U^\dagger_k G_{lm}].$$

$\psi^*\beta_m I_{kl}\psi$ is minus the complex conjugate of $\psi^*I_{kl}\beta_m\psi$. Therefore the spin term $S_{kl,m}$ is

$$\begin{aligned}S_{kl,m}&=-\frac{1}{2i}\psi^*(I_{kl}\beta_m+\beta_m I_{kl})\,\psi\\&=-\frac{1}{2\chi}\{G^{\dagger n}_k U_n g_{lm}-G^{\dagger n}_l U_n g_{km}-2G^\dagger_{km}U_l+2G^\dagger_{lm}U_k+\text{complex conjugate}\},\end{aligned} \tag{74}$$

so that for $\underline{k}$ and $\underline{l}$ not zero

$$S_{\underline{kl},0}=\frac{1}{\chi}[(G^\dagger_{\underline{k}0}U_{\underline{l}}-G^\dagger_{\underline{l}0}U_{\underline{k}})+\text{complex conjugate}].$$

The commutation rules for ψ and ψ^*, when the second quantization is introduced, have been given by Kemmer. They are‡

$$[(\psi^*\beta_0^2)_\mu,(\beta_0\psi')_\nu]=[(\psi^*\beta_0)_\mu,(\beta_0^2\psi')_\nu]=-(\beta_0^2)_{\nu\mu}\delta(x-x'), \tag{75}$$

where μ and ν refer to the matrix elements of the 1-row and 1-column matrices $\psi^*\beta_0^2$, $\psi^*\beta_0$ and $\beta_0\psi'$, $\beta_0^2\psi'$ respectively. Now

$$\begin{aligned}\psi^*\beta_0&=\psi^*(\beta\beta_1+\beta_0\beta)=U^{\dagger m}\Gamma^*_m\beta_0+\frac{i}{\chi}G^\dagger_{0m}\Gamma^{*m},\\ \beta_0\psi&=\beta_0 T_m U^m+\frac{i}{\chi}T^m G_{m0},\\ \psi^*\beta_0^2&=\psi^*(\beta\beta_0^2+\beta_0\beta\beta_0)=U^{\dagger m}(\Gamma_m g_{00}-\Gamma_0 g_{m0})+\frac{i}{\chi}G^\dagger_{0m}\Gamma^{*m}\beta_0\\&=U^{\dagger\underline{k}}\Gamma^*_{\underline{k}}+\frac{i}{\chi}G^\dagger_{0\underline{k}}\Gamma^{*\underline{k}}\beta_0,\\ \beta_0^2\psi&=T_{\underline{k}}U^{\underline{k}}+\frac{i}{\chi}\beta_0 T^{\underline{k}}G_{0\underline{k}},\end{aligned}$$

---

‡ In Kemmer's paper $[A,B]$ denotes a Poisson bracket, so that in his case

$$[A,B]=i(AB-BA).$$

Also $\delta_{\mu\nu}$ on the right is replaced by $(\beta_0^2)_{\nu\mu}$ here so that the right side vanishes automatically for the matrix elements corresponding to the eigenvalue 0 of β_0. This artifice of replacing $\delta_{\mu\nu}$ by $(\beta_0^2)_{\nu\mu}$ was pointed out to me by Professor Bhabha.

where the underlined index $\underline{k}$ runs from 1 to 3 only. From (75)

$$[(\psi^*\beta_0^2\mathcal{T}_{\underline{k}}), (\Gamma_{\underline{l}}^*\beta_0\psi')] = [(\psi^*\beta_0\mathcal{T}_{\underline{k}}), (\Gamma_{\underline{l}}^*\beta_0^2\psi')]$$
$$= -g_{\underline{kl}}\delta(x-x'),$$

i.e.
$$\left[U_{\underline{k}}^\dagger, \frac{i}{\chi}G'_{\underline{l}0}\right] = \left[\frac{i}{\chi}G_{0\underline{k}}^\dagger, U'_{\underline{l}}\right] = -g_{\underline{kl}}\delta(x-x'),$$

which are the correct commutation rules of the usual wave theory.

It is now extremely easy to include nuclear interaction in the above scheme. The total Lagrangian can be written as

$$L = L^H + L^M + L^E + L^I,$$

where L^H is the Lagrangian of the field of the heavy particle, L^M of the meson field and L^E of the electromagnetic field. L^H and L^M include the interaction of the electromagnetic field with the heavy particle and the meson respectively. L^I is the interaction term between the meson field and the heavy particles, and is given by

$$L^I = g_1\Psi^*\alpha^k(\tau_{PN}\Gamma_k^*\psi + \tau_{NP}\psi^*\mathcal{T}_k)\Psi + g_2\Psi^*\alpha^k\alpha^l(\tau_{PN}\Gamma_k^*\beta_l\psi + \tau_{NP}\psi^*\beta_l\mathcal{T}_k)\Psi. \quad (76)$$

Ψ and Ψ^* refer to the heavy particle. α^k's are the Dirac matrices acting on Ψ. $\Psi^* = \Psi^\dagger L$, where L is a matrix such that

$$\alpha^{\dagger k} = L\alpha^k L^{-1}.$$

τ_{PN} and τ_{NP} are the operators corresponding to the conversion of a neutron into a proton and vice versa. In fact (76) is exactly the same as the usual interaction term in the wave formulation of the meson theory. Obviously (76) can be extended to include pseudo-vector interaction by adding terms of the type‡

$$i\frac{g_1'}{3!}\Psi^*\alpha^k\alpha^l\alpha^m\epsilon_{klmn}(\tau_{PN}\Gamma^{*n}\psi + \tau_{NP}\psi^*\mathcal{T}^n)\Psi$$

and
$$\frac{g_2'}{2}\Psi^*\alpha^k\alpha^l\epsilon_{klmn}(\tau_{PN}\Gamma^{*m}\beta^n\psi + \tau_{NP}\psi^*\beta^n\mathcal{T}^m)\Psi,$$

where ϵ_{klmn} is a tensor antisymmetric in all the four indices and $\epsilon_{0123} = -1$.

Let us now consider the case of scalar mesons. Multiplying (1) by Γ^{*l} on the left

$$i\Gamma^{*l}\beta^k\partial_k\psi + \chi\Gamma^{*l}\psi = 0.$$

---

‡ If $\eta = 1-2\beta_0^2$ is taken as the reflexion matrix then $\psi^*\mathcal{T}_k$ is a vector and not a pseudo-vector, since

$$\eta\mathcal{T}_k = (1-2\beta_0^2)\mathcal{T}_k = \mathcal{T}_k - 2g_{00}\mathcal{T}_k + 2g_{0k}\mathcal{T}_0 = 2g_{0k}\mathcal{T}_0 - \mathcal{T}_k.$$

η satisfies the relations

$$\eta^2 = 1, \quad \eta\beta_k = -\beta_k\eta \quad (k \neq 0), \quad \eta\beta_0 = \beta_0\eta.$$

For the 5-row representation the reflexion matrix must be taken to be $-\eta$ so that $\psi^*\mathcal{T}_k$ may transform as a vector since in this case

$$\eta\mathcal{T}_k = \mathcal{T}_k - 2\mathcal{T}_0 g_{0k}.$$

Similarly, multiplying by $\Gamma^{*m}\beta_m$ and using (49b),

$$4i\partial_k \Gamma^{*k}\psi + \chi\Gamma^{*m}\beta_m\psi = 0.$$

Put
$$\Gamma^*_k\psi = U_k \tag{77a}$$

and
$$\Gamma^*_k\beta_l\psi = \tfrac{1}{4}g_{kl}\Gamma^*_m\beta^m\psi = -g_{kl}\frac{\chi}{i}U, \tag{77b}$$

then (1) is equivalent to

$$\partial_l U = U_l, \quad \partial_k U^k + \chi^2 U = 0,$$

which are again the usual equations for the scalar field. Also it is found that

$$\psi = (\beta + \tfrac{1}{4}\beta_k\beta\beta^k)\psi = T_m U^m - \tfrac{1}{4}\beta_k T^k \frac{\chi}{i} U$$

and
$$\beta^k\beta_l\psi = (\beta^k\beta\beta_l + \beta\beta^k\beta_l)\psi = -\frac{\chi}{4i}\delta^k_l\beta_m T^m U + T^k U_l,$$

so that (68) can be written as

$$T_k\partial_l U^k - \tfrac{1}{4}\beta_k T^k \frac{\chi}{i}\partial_l U = -\frac{\chi}{4i}\beta_k T^k\partial_l U + T^k\partial_k U_l$$

or
$$\partial_k U_l - \partial_l U_k = 0.$$

The expression for the current vector and the energy-momentum tensor can be transformed exactly as above. They agree with the usual expression of the wave formulation. Further, from (58) it is found that

$$\psi^*\beta_k\beta_l\beta_m\psi = \frac{\chi}{i}\{U^\dagger g_{kl}U_m - U^\dagger_k g_{lm}U\},$$

so that

$$S_{kl,m} = -\frac{1}{2i}\psi^*(I_{kl}\beta_m + \beta_m I_{kl})\psi = -\frac{\chi}{2}\{(U^\dagger_k g_{lm}U - U^\dagger_l g_{km}U) + \text{complex conjugate}\},$$

therefore $S_{\underline{kl},0} = 0$ in this case.

Also
$$\psi^*\beta_0 = \tfrac{1}{4}U^\dagger_0\Gamma^{*m}\beta_m + \frac{\chi}{i}U^\dagger\Gamma^*_0,$$

$$\psi^*\beta_0^2 = U^\dagger_0\Gamma^*_0 + \frac{\chi}{4i}U^\dagger\Gamma^{*m}\beta_m,$$

$$\beta_0\psi = \tfrac{1}{4}\beta_m T^m U_0 - \frac{\chi}{i}T_0 U,$$

$$\beta_0^2\psi = T_0 U_0 - \frac{\chi}{4i}\beta_m T^m U,$$

so that
$$\psi^*\beta_0 T_0 = \frac{\chi}{i}U^\dagger, \quad \psi^*\beta_0^2 T_0 = U^\dagger_0$$

and
$$\Gamma_0\beta_0\psi = -\frac{\chi}{i}U, \quad \Gamma_0\beta_0^2\psi = U_0,$$

therefore from (75)

$$\left[U_0^\dagger, -\frac{\chi}{i}U'\right] = \left[\frac{\chi}{i}U^\dagger, U_0'\right] = -\delta(x-x'),$$

which are the correct commutation rules of the usual scalar theory.

The nuclear interaction can be introduced in the scalar theory by exactly the same expression (76) as in the vector theory. However, on account of (49 *b*) and the fact that $\alpha^k\alpha_k = 4$ the second term in (76) can be written as

$$g_2\Psi^*(\tau_{PN}\Gamma_k^*\beta^k\psi + \tau_{NP}\psi^*\beta_k T^k)\,\Psi.$$

The g_2 term gives rise to the 'charge' interaction while the g_1 term brings about the 'dipole' interaction. This is precisely the reverse of what happens in the vector theory. The pseudo-scalar interaction can be introduced through the terms

$$\frac{ig_1'}{3!}\Psi^*\alpha^k\alpha^l\alpha^m\epsilon_{klmn}(\tau_{PN}\Gamma^{*n}\psi + \tau_{NP}\psi^*T^n)\,\Psi$$

and

$$\frac{ig_2'}{4!}\Psi^*\alpha^k\alpha^l\alpha^m\alpha^n\epsilon_{klmn}(\tau_{PN}\Gamma^{*r}\beta_r\psi + \tau_{NP}\psi^*\beta_r T^r)\,\Psi.$$

In the case of neutral mesons the additional condition

$$\psi^*T_k = \Gamma_k^*\psi \qquad (78a)$$

has to be imposed in both the scalar and the vector case. It then follows from the equations of motion that

$$\psi^*\beta_k T_l = -\Gamma_l^*\beta_k. \qquad (78b)$$

Also for this case τ_{NP} and τ_{PN} are both to be replaced by $\frac{1}{2}$.

The above method of treating neutral mesons is completely equivalent to the one employed by Majorana (1937) and Belinfante (1939). Let a bar denote the complex conjugate and a curl the transposed of a matrix. On taking the transposed of (2) and reversing the sign it is obvious that the matrices $-\tilde\beta_k$ satisfy the same commutation relations as β_k, so that for an irreducible representation of β_k there exists a matrix θ having the property that

$$-\tilde\beta_k = \theta\beta_k\theta^{-1}. \qquad (79a)$$

From (79 *a*) and the irreducibility of β_k it is easy to prove that (cf. Pauli 1936)

$$\tilde\theta = \pm\theta. \qquad (79b)$$

To decide between the two possibilities notice that from (11) and (79 *a*)

$$\tilde I^{kl}\theta + \theta I^{kl} = 0.$$

Put $$\Gamma_k = \tilde T_k.$$

Then $\Gamma_k\theta$ satisfies an equation similar to (14) and therefore

$$\Gamma_k\theta = a\Gamma_k^*,$$

where a is a number. Since θ is a non-singular matrix $a \neq 0$, and therefore without disturbing (79) θ can be replaced by θa^{-1}. In other words, θ can be so chosen consistently with (79) that

$$\Gamma_k \theta = \Gamma_k^*. \tag{79c}$$

On taking the transposed of (79c) and using (79b) it is found that

$$T_k = \pm \theta^{-1} \tilde{\Gamma}_k^*,$$

therefore $$\Gamma_k^* T_l = \pm \Gamma_k \theta . \theta^{-1} \tilde{\Gamma}_l^* = \pm \Gamma_k \tilde{\Gamma}_l^* = \pm \Gamma_l^* T_k = \pm g_{kl}.$$

It therefore follows on account of (28) that in (79b) only the positive sign is permissible. (78) can now be written as

$$\psi^* T_k = \Gamma_k^* \psi = \Gamma_k \theta \psi = \tilde{\psi} \theta T_k, \tag{80a}$$

$$\begin{aligned} \psi^* \beta_k T_l &= -\Gamma_l^* \beta_k \psi = -\Gamma_l \theta \beta_k \psi = \Gamma_l \tilde{\beta}_k \theta \psi \\ &= \tilde{\psi} \theta \beta_k T_l, \end{aligned} \tag{80b}$$

so that from (80) and (56), (57)

$$\psi^* = \tilde{\psi} \theta. \tag{81}$$

As a consequence of (81) the current density $\psi^* \beta_k \psi$ vanishes since

$$\psi^* \beta_k \psi = \tilde{\psi} \theta \beta_k \psi = \tilde{\psi} \tilde{\beta}_k \theta \psi = -\tilde{\psi} \theta \beta_k \psi.$$

Also the usual Lagrangian

$$L = \frac{1}{2i} (\partial_k \psi^* \beta^k \psi - \psi^* \beta^k \partial_k \psi) + \chi \psi^* \psi \tag{82a}$$

can be written in the form

$$L = \frac{1}{i} \psi \theta \beta^k \partial_k \psi + \chi \psi \theta \psi, \tag{82b}$$

where ψ is written instead of $\tilde{\psi}$ to emphasize the fact that now there is only one canonical variable ψ in the Lagrangian.

The connexion between θ and Λ is obtained from the equation

$$-\bar{\beta}_k = \theta^{\dagger -1} \Lambda \beta_k \Lambda^{-1} \theta^{\dagger} = \tilde{\Lambda}^{-1} \theta \beta_k \theta^{-1} \tilde{\Lambda},$$

so that $$\theta^{\dagger -1} \Lambda = c \tilde{\Lambda}^{-1} \theta,$$

where c is a number. Due to $\theta = \tilde{\theta}$ and $\Lambda = \Lambda^{\dagger}$ the above equation can be written as

$$\bar{\theta}^{-1} \Lambda = c \bar{\Lambda}^{-1} \theta.$$

Multiplying by $\bar{\Gamma}_k^*$ on the left

$$\bar{\Gamma}_k^* \bar{\theta}^{-1} \Lambda = c \bar{\Gamma}_k^* \bar{\Lambda}^{-1} \theta = c \Gamma_k \theta.$$

But $$\bar{\Gamma}_k^* \bar{\theta}^{-1} \Lambda = \bar{\Gamma}_k \Lambda = \Gamma_k^*,$$

therefore from (79c), $c = 1$ and

$$\bar{\theta}^{-1} \Lambda = \bar{\Lambda}^{-1} \theta. \tag{83a}$$

The matrix $\mathscr{L}$ of Belinfante (1939) corresponds to $\Lambda^{-1}\bar{\theta}$.

Let it also be remarked that $$\theta\beta = \tilde{\beta}\theta, \tag{83b}$$

since $$\theta\beta = \theta T_k \Gamma^k \theta = \bar{\beta}\theta.$$

5. Equation (1) can now easily be adapted for the case of zero rest-mass. Instead of (1) write

$$i\beta^k \partial_k \psi + \gamma\psi = 0, \tag{84}$$

where γ is a matrix satisfying the following conditions:

$$\gamma^2 = \gamma, \tag{85a}$$

$$\gamma\beta^k + \beta^k\gamma = \beta^k. \tag{85b}$$

In (84) γ appears in place of the mass χ of the particle. Multiplying (84) by $(1-\gamma)$ on the left one obtains

$$i\beta^k \partial_k(\gamma\psi) = 0. \tag{86}$$

Thus $\gamma\psi$ satisfies (1) with $\chi = 0$. Also on multiplying (84) by $\partial_m \beta^m \beta^l$ on the left it is found that

$$\partial_m \beta^m \beta^l \gamma\psi = \partial^l \gamma\psi, \tag{87}$$

therefore $\gamma\psi$ also satisfies (68). The second order wave equation

$$\partial_k \partial^k (\gamma\psi) = 0$$

follows from (86) and (87). Further, if ϕ is any solution of the equation

$$\beta^k \partial_k \phi = 0, \tag{88}$$

then $$\gamma\beta^k \partial_k \phi = \beta^k \partial_k (1-\gamma)\phi = 0.$$

Also since $$\gamma(1-\gamma)\phi = 0$$

$(1-\gamma)\phi$ satisfies (84). Thus $\psi + (1-\gamma)\phi$ is also a solution of (84). This corresponds to a gauge transformation of the usual theory.

Equation (84) can be derived from the Lagrangian

$$L = \frac{1}{i}[\partial_k \psi^* \beta^k \gamma\psi - \psi^* \gamma\beta^k \partial_k \psi] + \psi^* \gamma\psi. \tag{89}$$

γ must satisfy the condition $$\gamma^\dagger = \Lambda\gamma\Lambda^{-1}. \tag{90}$$

In fact from the result that for any irreducible representation of the β^k's there can exist only two γ's it follows that either (90) is fulfilled or

$$\Lambda\gamma\Lambda^{-1} = 1 - \gamma^\dagger.$$

In our case γ can be either β or $1-\beta$, and since

$$\Lambda\beta = \Lambda T_k \Gamma^{*k} = \Lambda T_k T^{\dagger k} \Lambda = \beta^\dagger \Lambda$$

(90) is satisfied. The Lagrangian (89) is gauge-invariant, i.e. it remains unchanged when ψ is replaced by $\psi+(1-\gamma)\phi$ and ϕ satisfies (88). This is why the expression (89) has been preferred to the following more obvious and simpler one,

$$L = \frac{1}{2i}\{\partial_k\psi^*\beta^k\psi - \psi^*\beta^k\partial_k\psi\} + \psi^*\gamma\psi,$$

which, though it is not gauge-invariant, yields the same equations of motion.

The expression for the current vector derived from (89) is the same as in the meson theory

$$S_k = \psi^*(\beta_k\gamma + \gamma\beta_k)\psi = \psi^*\beta_k\psi, \tag{91}$$

it is not gauge-invariant.

The energy-momentum tensor derived from (89) is

$$T_{kl} = \frac{1}{i}\{\partial_k\psi^*\beta_l\gamma\psi - \psi^*\gamma\beta_l\partial_k\psi\} + g_{kl}\psi^*\gamma\psi.$$

However, it can be verified that

$$T_{kl} = T'_{kl} - \frac{1}{2i}\partial^m\{\psi^*[(g_{kl}\beta_m - g_{km}\beta_l)\gamma - \gamma(g_{kl}\beta_m - g_{km}\beta_l)]\psi\}, \tag{92}$$

where
$$T'_{kl} = \frac{1}{2i}[\partial_k\psi^*\beta_l\psi - \psi^*\beta_l\partial_k\psi].$$

T'_{kl} has the same form as in the usual meson theory. Also it is found that due to (84)

$$T'_{kl} = \theta_{kl} - \frac{1}{2i}\partial^m[\psi^*(\beta_m\beta_k\beta_l - \beta_l\beta_k\beta_m)\psi],$$

where
$$\theta_{kl} = \psi^*(g_{kl} - \beta_k\beta_l - \beta_l\beta_k)\gamma\psi. \tag{93}$$

Therefore θ_{kl} which is symmetric in k and l may be taken to be the energy-momentum tensor instead of T_{kl}. (93) differs from (71) only in having γ instead of χ. Thus in the present theory γ formally plays a part somewhat similar to that of the mass in the meson theory. On account of the factor γ, θ_{kl} is gauge-invariant.

The expression (73) of the angular-momentum tensor remains valid also in the present case. However, the spin is not gauge-invariant.

The Hamiltonian formulation of (84) may be obtained in the same way as for the meson theory (Kemmer 1939). Multiplying (84) by $\beta_0(1-\gamma)$

$$\beta_0^2\partial_0(\gamma\psi) + \beta_0\beta_{\underline{k}}\partial^{\underline{k}}(\gamma\psi) = 0,$$

while from (87) with $l = 0$

$$(1-\beta_0^2)\,\partial_0(\gamma\psi) - \beta_{\underline{k}}\beta_0\,\partial^{\underline{k}}(\gamma\psi) = 0,$$

therefore
$$\partial_0(\gamma\psi) + (\beta_0\beta_{\underline{k}} - \beta_{\underline{k}}\beta_0)\,\partial^{\underline{k}}(\gamma\psi) = 0. \tag{94}$$

This gives the equation of motion of $\gamma\psi$. No such equation exists for $(1-\gamma)\psi$. This corresponds to the well-known fact that the Hamiltonian formulation gives the equations of motion only of the gauge-invariant quantities. Obviously $\gamma\psi$ is gauge-invariant while $(1-\gamma)\psi$ is not.

In order to eliminate the 'longitudinal waves' ψ is to be chosen in such a way that it satisfies in addition to (87) the equation

$$\partial_m \beta^m \beta^l \psi = \partial^l \psi. \tag{95}$$

This can be done by a suitable gauge transformation.

There are now two possibilities. Either $\gamma = 1-\beta$ or $\gamma = \beta$. Consider first the case of the 10-row representation of the β_k's so that now

$$\gamma = 1-\beta = \tfrac{1}{2}\beta_k \beta \beta^k.$$

It can be verified that in this case (84) corresponds to the usual Maxwell equations. The Lagrangian (89) becomes

$$\tfrac{1}{2}F^{\dagger kl}F_{kl} - (\partial_k U_l^\dagger F^{kl} + F^{\dagger kl}\partial_k U_l),$$

where
$$F_{kl} = \frac{1}{i}\Gamma_k^* \beta_l \psi \tag{96a}$$

and
$$U_k = \Gamma_k^* \psi. \tag{96b}$$

From (69) it follows that the elimination of the longitudinal waves in this case corresponds to imposing the condition $\partial^k U_k = 0$. With this choice of γ (84) may be used to describe the photon or the electromagnetic field. In this case the reality condition (78) or (81) has to be imposed. The interaction of the photon with the charged particles can be introduced in a way entirely similar to (76). For example, to introduce interaction with an electron of charge e the interaction term

$$e\Psi^* \alpha^k \Psi \Gamma_k \theta \psi$$

has to be added to the sum of the independent Lagrangians of the electromagnetic and electron fields. On account of (81) and (83), (89) can now be written in the form

$$L' \equiv \tfrac{1}{2}L = i\psi\theta\gamma\beta^k \partial_k \psi + \tfrac{1}{2}\psi\theta\gamma\psi. \tag{97}$$

For the other choice $\gamma = \beta$ (84) is equivalent to

$$\partial_k F^{lk} = U^l, \tag{98a}$$

$$\partial_k U_l - \partial_l U_k = 0, \tag{98b}$$

F^{kl} and U_l being still defined by (96). It follows from (98) that there exists a U such that
$$U_l = \partial_l U \quad \text{and} \quad \partial^l \partial_l U = \partial^l U_l = 0.$$

Equations (98) are thus formally equivalent to the equations of the usual scalar theory. The energy-monentum tensor turns out to be

$$U_k^\dagger U_l + U_l^\dagger U_k - g_{kl} U^{\dagger m} U_m,$$

and therefore agrees with the usual expression. The current vector and the commutation rules for the second quantization are, however, different. A gauge transformation consists in adding to F^{kl} another antisymmetrical tensor F'^{kl} satisfying

$$\partial^k F'_{kl} = 0.$$

The elimination of the longitudinal waves corresponds to imposing the condition

$$\partial_k F_{lm} + \partial_m F_{kl} + \partial_l F_{mk} = 0.$$

Coming to the 5-row representation of the β_k's and taking $\gamma = 1-\beta = \frac{1}{4}\beta_k \beta \beta^k$, it is found that (78) is equivalent to

$$\partial_m U^m + U = 0, \tag{99a}$$

$$\partial_k U = 0, \tag{99b}$$

where U and U_m are defined by (77) with $\chi = 1$. The energy-momentum tensor comes out to be

$$-\theta_{kl} = g_{kl} U^\dagger U.$$

Gauge transformation consists in adding to U_m a tensor U'_m such that

$$\partial^m U'_m = 0.$$

(88) corresponds to imposing the condition $\partial_k U_l = \partial_l U_k$. The current vector has the same form as in the usual scalar theory

$$\frac{1}{i}\{U^\dagger U_k - U_k^\dagger U\}.$$

Since from (99) U is a constant the energy-momentum density is also constant and therefore the field is of no physical interest.

The case $\gamma = \beta$ corresponds to the usual scalar theory, the equations being

$$\partial^k U = U^k, \quad \partial_k U^k = 0.$$

The Lagrangian is $\qquad U_k^\dagger U^k - (\partial^k U^\dagger . U_k + U_k^\dagger \partial^k U).$

Gauge transformation consists in adding to U^k a vector U'^k satisfying

$$\partial_k U'^k = 0.$$

In this case (95) follows automatically from (87) since

$$\begin{aligned} \partial_k \beta^k \beta^l \beta \psi - \partial^l \beta \psi &= \partial_k \beta^k (1-\beta) \beta^l \psi - \partial^l \beta \psi \\ &= \partial_k \beta^k \beta^l \psi - \partial_k g^{kl} (1-\beta) \psi - \partial^l \beta \psi \quad \text{from (57b)} \\ &= \partial_k \beta^k \beta^l \psi - \partial^l \psi. \end{aligned}$$

6. Till now the theory has been developed in an abstract way. It is useful to know what the matrices T_k, Γ^{*k} and β actually are for a particular representation. For this purpose it is best to choose the following representation. For the 10-row representation, ψ and $\beta_{\alpha\dot{\lambda}} = \beta_k \sigma^k_{\alpha\dot{\lambda}}$ are

$$\psi = \begin{pmatrix} \psi(0,1) \\ \psi(\frac{1}{2},\frac{1}{2}) \\ \psi(1,0) \end{pmatrix}, \quad \beta_{\alpha\dot{\lambda}} = \sqrt{2} \begin{pmatrix} 0 & v_\alpha u^+ & 0 \\ -\frac{1}{2} u_\alpha v_{\dot{\lambda}}^+ & 0 & -\frac{1}{2} v_\alpha^+ u_{\dot{\lambda}} \\ 0 & u_\alpha^+ v_{\dot{\lambda}} & 0 \end{pmatrix}, \tag{100}$$

where $v_\alpha = v_\alpha(\frac{1}{2})$ and $v_\alpha^+ = v_\alpha(1)$ and similarly for others. The splitting up of ψ corresponds to the decomposition of the whole representation with respect to the k and l values so that the components of ψ are $\psi(k, l)$ (cf. Bhabha 1945). For this representation

$$\frac{1}{\sqrt{2}} T_{\alpha\dot{\lambda}} = \begin{pmatrix} 0 \\ u_\alpha u_{\dot{\lambda}} \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \Gamma^*_{\alpha\dot{\lambda}} = \overgroup{\undergroup{0 \quad v_\alpha v_{\dot{\lambda}} \quad 0}}, \quad \beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad (101a)$$

Also

$$\omega = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (101b)$$

(101b) is easily obtained from (59b), (100) and (101a) if one notices that the spinor corresponding to ϵ^{klmn} is given by (cf. Harish-Chandra 1945)

$$\tfrac{1}{2}\epsilon^{klmn}\sigma_{k,\alpha\dot{\lambda}}\sigma_{l,\beta\dot{\mu}}\sigma_{m,\gamma\dot{\nu}}\sigma_{n,\delta\dot{\rho}} = \epsilon_{\alpha\beta}\epsilon_{\gamma\delta}\dot{\Sigma}_{(\dot{\lambda}\dot{\mu})(\dot{\nu}\dot{\rho})} + \epsilon_{\dot{\lambda}\dot{\mu}}\epsilon_{\dot{\nu}\dot{\rho}}\Sigma_{(\alpha\beta)(\gamma\delta)}, \qquad (102a)$$

where and

$$\left.\begin{aligned} \Sigma_{(\alpha\beta)(\gamma\delta)} &= i(\epsilon_{\alpha\gamma}\epsilon_{\beta\delta} + \epsilon_{\beta\gamma}\epsilon_{\alpha\delta}) \\ \dot{\Sigma}_{(\dot{\lambda}\dot{\mu})(\dot{\nu}\dot{\rho})} &= -i(\epsilon_{\dot{\lambda}\dot{\nu}}\epsilon_{\dot{\mu}\dot{\rho}} + \epsilon_{\dot{\mu}\dot{\nu}}\epsilon_{\dot{\lambda}\dot{\rho}}). \end{aligned}\right\} \qquad (102b)$$

On the other hand, for the 5-row representation

$$\psi = \begin{pmatrix} \psi(\frac{1}{2}, \frac{1}{2}) \\ \psi(0, 0) \end{pmatrix}, \quad \beta_{\alpha\dot{\lambda}} = \sqrt{2}\begin{pmatrix} 0 & u_\alpha u_{\dot{\lambda}} \\ v_\alpha v_{\dot{\lambda}} & 0 \end{pmatrix},$$

$$\frac{1}{\sqrt{2}} T_{\alpha\dot{\lambda}} = \begin{pmatrix} u_\alpha u_{\dot{\lambda}} \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \Gamma^*_{\alpha\dot{\lambda}} = \overgroup{\undergroup{v_\alpha v_{\dot{\lambda}} \quad 0}}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

REFERENCES

Belinfante 1939 *Physica*, **6**, 849–887.
Bhabha 1945 *Rev. Mod. Phys.* **17**, 200–216.
Booth & Wilson 1940 *Proc. Roy. Soc.* A, **175**, 483–518.
Dirac 1936 *Proc. Roy. Soc.* A, **155**, 447–459.
Fierz 1939 *Helv. Phys. Acta*, **12**, 3–37.
Harish-Chandra 1945 *Proc. Ind. Acad. Sci.* **23**, 152–163.
Kemmer 1939 *Proc. Roy. Soc.* A, **173**, 91–116.
Kemmer 1943 *Proc. Camb. Phil. Soc.* **39**, 189–196.
Majorana 1937 *Nuovo Cim.* **14**, 171.
Pauli 1936 *Ann. Inst. Poincaré*, **6**, 109–136.
Schrödinger 1943 *Proc. Roy. Irish Acad.* A, **48**, 135.
Wilson 1940 *Proc. Camb. Phil. Soc.* **36**, 363–380.

INDEX TO VOLUME 186 (A)

Airwave surrounding an expanding sphere (Taylor), 273.
Axisymmetric turbulence, theory (Batchelor), 480.

Bakerian Lecture (Haworth), 1.
Balmer spectrum. Recombination and long duration (Zanstra), 236.
Barrell, H. The standards of length in wave-lengths of light (discussion), 164.
Batchelor, G. K. The theory of axisymmetric turbulence, 480.
Bell, R. P., Everett, D. H. & Longuet-Higgins, H. C. Kinetics of the base-catalysed bromination of diethyl malonate, 443.
Boas, W. & Honeycombe, R. W. K. The plastic deformation of non-cubic metals by heating and cooling, 57.
Bolland, J. L. Kinetic studies in the chemistry of rubber and related materials. I. The thermal oxidation of ethyl linoleate, 218.
Bradley, R. S., Evans, M. G. & Whytlaw-Gray, R. The rate of evaporation of droplets. Evaporation and diffusion co-efficients and vapour pressures of dibutyl phthalate and butyl stearate, 368.

Chromatography, contributions to the theory (Glückauf), 35.
Clark, J. S. The acceleration due to gravity (discussion), 192.
Conservation of momentum (Milne), 432.
Contributions to the theory of chromatography (Glückauf), 35.
Copson, E. T. An integral equation method of solving plane diffraction problems, 100.
Correspondence between the particle and the wave aspects of the meson and the photon (Harish-Chandra), 502.
Craggs, J. D. & Meek, J. M. The emission of light from spark discharges, 241.
Cullis, C. F. & Hinshelwood, C. N. Mechanism of the hydrogen-oxygen reaction, IV and V, 462.

Dadson, R. S. Acoustical standards of measurement (discussion), 207.
Darwin, Sir Charles & others. Discussion on Units and Standards, 149.
Diethyl malonate. Kinetics of the base-catalysed bromination (Bell, Everett and Longuet-Higgins), 443.
Diffraction of sound pulses, I, II, III, IV (Friedlander), 322.
Discussion on Units and Standards (Darwin & others), 149.
Divergence difficulty of quantized field theories and the rigorous treatment of radiation reaction (Peng), 119.
Drag of cylinders in fluids at slow speeds (White), 472.

Emission of light from spark discharges (Craggs & Meek), 241.
Evans, M. G. *See* Bradley, Evans & Whytlaw-Gray.
Everett, D. H. *See* Bell, Everett & Longuet-Higgins.
Experimental production of red and yellow sandstones from chalybeate water (Rayleigh), 411.
Explosion waves and shock waves (Payman & Shepherd), 293.

Fluid motion between parallel planes. Dynamical stability (Meksyn), 391.
Foucault Test, contribution to the theory (Linfoot), 72.

Friedlander, F. G. The diffraction of sound pulses, I, II, III, IV, 322.
Frost, R. Turbulence and diffusion in the lower atmosphere, 20.

Gas-molecules, probability of encounters (Richardson), 422.
Glückauf, E. Contributions to the theory of chromatography, 35.
Gould, F. A. The barometric standard (discussion), 195.
Gould, F. A. The standards of mass (discussion), 171.

Hall, J. A. The temperature scale (discussion), 179.
Harish-Chandra. The correspondence between the particle and the wave aspects of the meson and the photon, 502.
Hartshorn, L. Electrical standards (discussion), 185.
Haworth, W. N. The structure, function and synthesis of polysaccharides (Bakerian Lecture), 1.
Hinshelwood, C. N. *See* Cullis & Hinshelwood.
Honeycombe, R. W. K. *See* Boas & Honeycombe.

Integral equation method of solving plane diffraction problems (Copson), 100.
Invariant form for the prior probability in estimation problems (Jeffreys), 453.

Jeffreys, H. An invariant form for the prior probability in estimation problems, 453.

Kinetic studies in the chemistry of rubber and related materials. I. The thermal oxidation of ethyl linoleate (Bolland), 218.

Light Division. The photometric units and standards (discussion), 204.
Linfoot, E. H. A contribution to the theory of the Foucault Test, 72.
Longuet-Higgins. *See* Bell, Everett & Longuet-Higgins.

Mechanism of the hydrogen-oxygen reaction, IV and V (Cullis & Hinshelwood), 462.
Meek, J. M. *See* Craggs & Meek.
Meksyn, D. Fluid motion between parallel planes. Dynamical stability, 391.
Mica, further interferometric studies. New multiple beam fringes and their application (Tolansky), 261.
Milne, E. A. On the conservation of momentum, 432.

Payman, W. & Shepherd, W. C. F. Explosion waves and shock waves, 293.
Peng, H. W. On the divergence difficulty of quantized field theories and the rigorous treatment of radiation reaction, 119.
Plastic deformation of non-cubic metals by heating and cooling (Boas & Honeycombe), 57.
Polysaccharides, structure, function and synthesis (Haworth, Bakerian Lecture), 1.

Rate of evaporation of droplets. Evaporation and diffusion co-efficients and vapour pressures of dibutyl phthalate and butyl stearate (Bradley, Evans & Whytlaw-Gray), 368.
Rayleigh, Lord. Experimental production of red and yellow sandstones from chalybeate water, 411.
Richardson, L. F. The probability of encounters between gas-molecules, 422.

Sears, J. E. The standards of length (discussion), 152.
Shepherd, W. C. F. *See* Payman & Shepherd.
Standards, Units and (discussion), 149.
Stott, V. The measurement of volume (discussion), 200.

Taylor, Sir Geoffrey. The airwave surrounding an expanding sphere, 273.
Tolansky, S. Further interferometric studies with mica. New multiple beam fringes and their application, 261.
Turbulence and diffusion in the lower atmosphere (Frost), 20.

Units and Standards (discussion), 149.

White, C. M. The drag of cylinders in fluids at slow speeds, 472.
Whytlaw-Gray, R. *See* Bradley, Evans & Whytlaw-Gray.

Zanstra, H. Recombination and the long duration of the Balmer spectrum, 236.

END OF THE ONE HUNDRED AND EIGHTY-SIXTH VOLUME (SERIES A)

I.A.R.I. 75

INDIAN AGRICULTURAL RESEARCH
INSTITUTE LIBRARY, NEW DELHI.

Date of Issue	Date of Issue	Date of Issue
28/3/58		

GIPNLK—H-40 I.A.R.I.—29-4-5—15,000